

CARBON ATMOSPHERIC TRACER RESEARCH TO IMPROVE NUMERICAL SCHEMES AND EVALUATION



CATRINE

Carbon Atmospheric Tracer
Research to Improve
Numerics and Evaluation

D1.1 Evaluation and analysis of IFS tracer advection scheme accuracy and conservation properties and its ability to estimate emissions

Due date of deliverable	31 March 2025
Submission date	25 March 2025
File Name	CATRINE-D1-1-V1.1
Work Package /Task	WP1/ Task 1.1
Organisation Responsible of Deliverable	ECMWF
Author name(s)	Michail Diamantakis, Sylvie Malardel, Luca Cantarello, Claire Laurent, Giovanni Tumolo, Stefan Versick and Anna Agusti-Panareda
Revision number	1.1
Status	Issued
Dissemination Level / location	Public



Funded by the
European Union

The CATRINE project (grant agreement No 101135000) is funded by the European Union.

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

The advection scheme is a numerical method used to solve the Partial Differential Equations (PDEs) which represent the resolved (non-parametrized) part of the transport. In this report, we present numerical modelling aspects related to the IFS advection scheme and their impact on its ability to accurately transport atmospheric tracers and estimate emissions from various sources. The main goal of this deliverable is to evaluate the core global CO2MVS model (IFS) advection schemes in terms of mass conservation properties and their effect on the accuracy of tracer gas simulations. This evaluation is motivated by the fact that the IFS is not inherently mass conserving; instead, global conservation is enforced using additional mass fixer algorithms that employ a weighted approach based on local error estimates.

To understand the origin and evolution of tracer mass conservation errors in the IFS advection scheme, we begin with idealized case studies, designed to differentiate conservation errors from transport errors and to subject the scheme to extreme conditions. This allows us to rigorously evaluate the effectiveness of current and proposed mass conservation solutions, which will be detailed in CATRINE deliverable D1.2. Building upon these fundamental analyses, we progress to complex, full-scale greenhouse gas simulations, adhering to CATRINE's WP5 and WP7 protocols. These simulations incorporate sophisticated emission protocols, enabling us to assess the impact of different algorithms on mass conservation and tracer accuracy. By comparing these realistic simulations with observational data, we gain valuable insights for further refinement of the IFS advection scheme. Ultimately, this comprehensive approach guides the development of improved algorithms for enhanced mass conservation and tracer simulation accuracy.

This work is closely linked to other work packages, contributing to the setup of inter-comparison studies. Additionally, we document our progress on initial sensitivity experiments using the newly developed Ensemble Data Assimilation based idealised inversion system (Observing System Simulation Inversion Experiment), which incorporates synthetic observations. The findings from this deliverable have provided valuable guidance to steer developments which will take place in WP2. A key outcome of this exercise is the demonstration that, despite the IFS lacking inherent mass conservation, it can still accurately simulate greenhouse gas transport with the use of mass fixer algorithms. Thanks to its computationally efficient design, this can be easily achieved at resolutions up to 4.5km. We anticipate that further improvements in the IFS transport accuracy will be made available by the end of the CATRINE project.

Table of Contents

1	Executive Summary	2
2	Introduction	4
2.1	Background	4
2.2	Scope of this deliverable.....	5
2.2.1	Objectives of this deliverable	5
2.2.2	Work performed in this deliverable.....	6
2.2.3	Deviations and counter measures.....	6
2.3	Project partners:	6
3	A description of the IFS tracer advection scheme and its properties	7
4	Sensitivity tests from idealised cases	9
4.1	Solid body rotation.....	9
4.2	Single column rising bubble case.....	12
4.3	Advection over a mountain range	15
5	Advection sensitivity tests with plume type tracers in real weather scenarios.....	19
5.1	Large plume tracers without emissions	19
5.2	Plume tracers emitted by a single-point source.....	23
5.3	The impact of the lower boundary.....	25
5.4	The impact of resolution in mass conservation error.....	26
6	Advection sensitivity tests on CATRINE simulations using inter-comparison protocols	28
7	ICON-ART and comparisons with the IFS	35
7.1	Tests for ICON-ART transport schemes	35
7.1.1	Preliminary results from artificial plume type tracers	38
7.1.2	Preliminary results from real world tracers	39
8	Sensitivity in the IFS EDA-based Observing System Simulation Inversion Experiments (OSSE)	39
8.1	Synthetic emission inversion experiments	40
8.2	Experiment design and results.....	40
8.3	Discussion and plans for further experiments	44
9	Concluding remarks and future work.....	45
10	References	48

2 Introduction

2.1 Background

The ECMWF Integrated Forecast System (IFS) is a multi-purpose, highly efficient, and accurate global model for weather forecasting and atmospheric composition. Combined with a 4-dimensional variational data assimilation system (4D-VAR, Rabier et al., 2000), it generates accurate atmospheric analyses every six hours, providing initial conditions for global and regional models. The IFS also underpins the ERA5 system which produces high-quality meteorological reanalysis datasets which are widely used by organizations worldwide. Its sophisticated technical and scientific infrastructure, its Numerical Weather Prediction (NWP) accuracy (the most accurate global model in the world according to WMO metrics) and its advanced atmospheric composition capabilities make it an ideal choice for a 4D-VAR atmospheric inversion system. Furthermore, the IFS benefits from collaborative developments across European member states. For example, it is worth mentioning the strong collaboration with Meteo-France and the ACCORD (A Consortium for CONvection-scale modelling Research and Development) consortium on atmospheric dynamics. In fact, the ECMWF IFS and the global model ARPEGE of Meteo-France use the same dynamical core.

The IFS dynamical core, the part of the model which solves the governing Partial Differential Equations (PDEs) of atmospheric motion, relies on a spectral transform method with a semi-implicit semi-Lagrangian (SISL) time-stepping scheme. The semi-implicit (SI) part deals with the integration of various forcing (source) terms while the semi-Lagrangian (SL) advection scheme computes the resolved part of the transport of the air and its different constituents. The unconditional stability and good dispersion properties of the SL advection scheme, combined with the unconditional stability properties of the SI scheme, enable accurate and stable long time-step integrations (not limited by the CFL number) with small phase errors. SL advection is not only efficient because it allows long time-steps, but it is also multi-tracer efficient which means that many tracers can be transported with a relatively small increase of computational cost compared with Eulerian advection schemes. These are excellent properties for an atmospheric composition system and an atmospheric inversion system.

Important properties for an advection scheme in a transport model are good accuracy, monotonicity, positive definiteness and mass conservation for air and its constituents (tracers). A semi-Lagrangian (SL) advection scheme can satisfy all these properties except of the last one. It can preserve the global mass of air and tracers only under very specific conditions. For instance, in a uniform, non-divergent flow on a regular grid, far from surface boundaries, and with a smooth tracer field, SL advection without use of additional filters for positivity and monotonicity can be conservative in each flow direction. However, such idealised conditions are rarely encountered in practice and the global mass of a tracer or the mass of air after advection differs from its global mass before advection. This difference is known as the “global mass conservation error”. Until recently, in global NWP, this error was less critical due to the smoothness of fields, the strong effect of parametrizations, and the fact that each assimilation cycle disrupts conservation. The situation is different for atmospheric composition modelling, climate simulations, and high-resolution weather forecasts (e.g. cloud-resolving models) where maintaining mass conservation is critical. Artificial mass growth due to conservation error can impact the simulated climate or produce wrong estimates of fluxes in an atmospheric inversion system and therefore it must be controlled.

To correct the mass conservation error, the IFS incorporates mass fixer algorithms that adjust tracer concentrations to ensure global mass conservation after advection [Diamantakis and Flemming, 2014]. The complexity of these algorithms has evolved over

CATRINE

time in IFS, from the two-dimensional idealised advection case of Bermejo and Conde [2002] to the fully three-dimensional mass fixer for complex atmospheric flows described in Diamantakis and Agusti-Panareda [2017]. This method enforces global mass conservation while preserving essential transport properties such as monotonicity and positive definiteness, minimizing the impact on local accuracy. Previous studies have shown that the IFS mass fixer improves accuracy in greenhouse gas simulations [Agusti-Panareda et al., 2017] and enhances energy balances and precipitation forecasts when applied to moist tracers [Becker et al., 2022]. However, a drawback is that these corrections are applied independently of the governing equations, adjusting tracer fields, air density, or surface pressure through external criteria. A desirable goal—aligned with the objectives of the CATRINE project—is to reduce conservation errors so that only minimal corrections are needed at each time step.

It is important to emphasize that mass conservation error is only one aspect of advection or overall transport error. The latter is defined as the difference of the numerical model simulation (solution) from the truth, which may be represented by observations, an atmospheric analysis or an analytical solution in simpler idealised cases. Standard statistical measures such as bias, Root Mean Square Error (RMSE), standard deviation of error, Mean Absolute Error (MAE), anomaly correlation coefficient can be applied to assess transport errors. An accurate advection scheme should transport fields at the correct speed and preserve sharp gradients and smooth fields without introducing excessive numerical diffusion or dispersion. However, some numerical schemes may conserve mass but lack accuracy. For example, a first-order upwind finite-volume scheme is conservative by construction, but highly dispersive, resulting in reduced accuracy.

The transport error also encompasses errors from parametrized unresolved transport processes, such as convection and turbulence, which will be examined in WP5-6 of CATRINE. While these errors are likely to be larger than the advection mass conservation error at global scale, a significant and systematic uncorrected conservation error can still contribute to the overall transport error and degrade the simulation accuracy. Finally, as demonstrated by Eastham and Jacob [2017], lack of resolution can lead to numerical diffusion contributing to the overall transport error. The IFS is well-suited in this regard, as it can easily handle horizontal resolutions up to 4.5km with 137 vertical levels.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverable

The primary objectives of this deliverable are: (i) to evaluate the advection scheme of the global CO2MVS model IFS in terms of mass conservation properties and transport accuracy relevant for its ability to estimate emissions (ii) establish how different algorithms contribute to the conservation errors and the quality of the transport scheme, and (iii) identify areas for improvement providing guidance for model developments currently in progress and which will be completed in WP2. Emphasis will be placed on the sensitivity of mass conservation and advection scheme errors with respect to fundamental components of the IFS advection scheme such as numerical algorithms used for interpolation, positive definiteness and shape preservation (monotonicity). The impact of the IFS tracer mass fixer in local accuracy will be analysed while the overall quality of the IFS advection scheme will be assessed through idealised academic, simplified and real greenhouse tracer simulation case studies following protocols defined in WP5 and WP7.

2.2.2 Work performed in this deliverable

In the following section (3) of this report, we describe briefly how the IFS advection scheme works, its main components and their purpose. This helps us to identify areas which are suitable for conducting sensitivity experiments to understand the origins of mass conservation error and the impact of various advection scheme algorithms on conservation and accuracy. In Section 4, we describe idealised dynamics case studies. We start progressively from horizontal advection cases to vertical advection dominated cases and progress to full three-dimensional cases examining the impact on conservation and accuracy either on a precise (when the solution is known) or a qualitative way (when the exact solution or truth is not available). In Section 5, we step up the complexity of testing using idealised plume-type tracers in a real weather setup (i.e. winds, pressure and remaining prognostic variables are produced from an IFS weather forecast). In Section 6, we present advection sensitivity tests using accurate inter-comparison protocols for fluxes which have been developed in WP7 and WP5. This enables comparisons against a range of observations and gives valuable feedback for the performance of the IFS model and its sensitivity to advection helping to focus on a future optimal tuning. In Section 7, we present preliminary comparisons between IFS and ICON-ART model on two types of cases: idealised plume-type transported in a weather forecast setup and results from comparing the two models using WP5 protocols. In Section 8, we finalise this work presenting results from testing an Ensemble Data Assimilation (EDA) based Observing System Simulation Inversion system which has been developed recently by researchers working in CAMS section of ECMWF. In Section 9 the report concludes presenting a summary of our findings which gives feedback on where the focus of developments for improving the advection scheme should be.

It is worth emphasizing that the cases presented in Sections 4 and 5, are intentionally extreme, designed to stress test the SL advection scheme and its components such as mass fixers and monotone limiters. Consequently, the mass conservation errors observed in these cases are much larger than those encountered in the realistic scenario greenhouse tracer simulations presented in Section 6. In realistic greenhouse simulations, mass conservation error growth tends to be smaller due to the influence of physical processes such as the parametrized turbulent mixing, which tend to smooth tracer gradients. The gradients themselves are also weaker because of substantial tracer background values and weaker emissions. However, these idealized scenarios serve as valuable diagnostic tools, helping us to expose numerical modelling issues, understand the limits of our algorithms, and refine our schemes for CO2MVS and for other atmospheric composition cases. Such refinements will be detailed in Deliverable D1.2.

2.2.3 Deviations and counter measures

none

2.3 Project partners:

Partners	
EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS	ECMWF
COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES	CEA
METEO-FRANCE	METEO-FRANCE
WAGENINGEN UNIVERSITY	WU
KARLSRUHER INSTITUT FUER TECHNOLOGIE	KIT
HELSINGIN YLIOPISTO	UH
UNIVERSITE DE REIMS CHAMPAGNE-ARDENNE	URCA
ALBERT-LUDWIGS-UNIVERSITAET FREIBURG	UFR

3 A description of the IFS tracer advection scheme and its properties

To define SL advection in the IFS we consider a tracer with specific ratio (against moist air) $\phi = \rho_\phi/\rho$. The tracer advection problem in the so-called non-conservation form which is used by the IFS is defined by the equation:

$$\frac{D\phi}{Dt} = 0, \quad (1)$$

which is equivalent, and actually can be derived from, the following two PDEs written in “conservative” or “flux” form:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{V}) = 0 \quad (2)$$

$$\frac{\partial \rho_\phi}{\partial t} + \nabla \cdot (\rho_\phi \vec{V}) = 0. \quad (3)$$

In (1) the advected variable ϕ is a scalar quantity, the mixing or specific mixing ratio measured in (kg of tracer)/(kg of air). We will often use the term “concentration” to denote this quantity. D is the “material” or “Lagrangian” derivative operator which is defined as $D\phi/Dt = \partial\phi/\partial t + \vec{V} \cdot \nabla\phi$, where ∇ is the 3D nabla (gradient) operator applied in spherical coordinates and \vec{V} is the 3D wind that advects the tracer ϕ . In (3), the advected variable is the tracer density $\rho_\phi = \rho\phi$. The IFS SL advection scheme solves (1) by first finding the upstream trajectory points of model grid points which are called the “departure points”. This is an important calculation as insufficient accuracy in the calculation of the departure points can introduce phase errors which will be manifested as errors in all prognostic fields including tracer concentrations. In IFS this is done accurately using a scheme called SETTLS where its details can be found in Hortal [2002] and Diamantakis and Vana [2022]. The grid used in IFS is the so called cubic-octahedral grid [Malardel et al., 2016], a quasi-uniform grid with approximately equal spacing at different latitudes which allows accurate transformation of prognostic variables from spectral to grid-point space and vice-versa. Once departure points have been computed, the numerical solution of the advection problem is found:

$$\phi^{t+\Delta t} = \phi^{t_D}: \quad \text{Find } D \text{ locations} \rightarrow \text{Interpolate } \phi \text{ at } D \text{ locations}: \quad \phi^{t_D} = \phi(r_D, t).$$

where r_D gives the coordinates of the departure points. The transported field is mapped onto the departure points using a Lagrange type interpolation method in three-dimensional space. The global mass of the tracer is the surface integral of the total column mass of a tracer and in discrete form is defined [see Diamantakis and Flemming, 2014]:

$$M = \sum_{j=1}^N A_j \sum_{k=1}^K \phi_{jk} \frac{\Delta p_{jk}}{g} \quad (4)$$

where g the acceleration of gravity, K is the number of vertical levels, N the number of grid-points at each level, A_j is the surface area of an assumed grid-box around grid-point j (constant in a vertical column) and Δp_{jk} is the pressure thickness of the layer around grid-point jk . The dependence of the tracer mass on surface pressure is an important factor which implies that the conservation error of a tracer also depends on ambient air density (or equivalently pressure). It is straightforward to derive the total mass of air setting $\phi=1$ in equation (4). Henceforth, when we compute the conservation error, we will

CATRINE

apply equation (4) with the corresponding ϕ , Δp values before and after advection. We will often express this error as a “percent of initial mass” i.e.

$$E_{mass} = 100 \times \frac{M_{adv} - M_0}{M_0} \quad (5)$$

where M_0 , M_{adv} is the tracer mass before and after advection.

While the SL advection has very small phase errors, the interpolation step can introduce amplitude errors which depend on the smoothness of the advected field and the interpolation method order. Interpolation errors are a big contributing factor to global mass conservation errors. In all operational configurations of the IFS, such errors are corrected using the “a-posteriori” mass fixers mentioned in the introduction. It is worth emphasizing that, with the cubic grid [Malardel et al., 2016], the global mass conservation error for the air is very small: less than 0.001% after 10 days of forecast. A distinction must be made between tracer mass error and the total atmospheric air mass error. In the latter, a special treatment for the surface pressure advection helps to significantly reduce such errors. Furthermore, the higher accuracy in total air mass conservation, is also a reflection of the overall accuracy of the spectral SISL formulation: the continuity equation links all prognostic dynamical variables, and an accurate conservation of the air reflects this accuracy. In the following sections, the problem of tracer mass conservation in IFS will be thoroughly analysed showing that tracer advection and its associated conservation error is particularly sensitive to interpolation errors and interpolation scheme choices, the boundary conditions, the direction of the transport (horizontal versus vertical), the type of the flow (divergent/convergent or divergent free), the atmospheric layer, and the shape/type of tracer.

With a flux-form formulation such as (2), (3), it is easy to define the concept of “local conservation”: the mass that leaves a discrete volume cell will enter its neighbouring cell and the net balance is zero. A finite-volume scheme based on a flux-form formulation is then conservative by construction. However, with a SL method such as the one described by (1), because of its Lagrangian and pointwise approach instead of a volume element approach, it is not possible to diagnose local conservation errors and therefore to enforce local conservation. To circumvent this problem, we developed a range of tests with highly localised tracers such as plume tracers, bubble cases and point source emissions in an initially “empty” atmosphere. With such tracers, because of their locality, in relatively short lead times, we can infer conclusions on local mass as well as global mass conservation errors and how these grow in the horizontal and vertical directions.

As mentioned earlier, the IFS SL advection scheme relies on interpolation at the departure points. The interpolation method used in the IFS for tracers is a three-dimensional cubic-Lagrange method. This is enhanced with interpolation limiters which enforce monotonicity in advection i.e. they do not allow generation of new minima and maxima when a tracer field is transported by the dynamics passively in the absence of source/sinks or other physical process which may change its value. A variant of the IFS cubic-Lagrange interpolation method is the COMAD method described in Malardel and Ricard [2015]. The same limiters applied in standard cubic-Lagrange interpolation can be applied in the COMAD scheme. Furthermore, in IFS a linear interpolation is also available, which although in practice yields very diffusive results and poor forecast accuracy, is quite valuable for investigations because of its simplicity and because of its inherent monotonicity it does not require interpolation limiters.

In the discussion that follows we are going to use the following conventions: (i) we will refer to simulations with COMAD scheme as SL-COMAD or COMAD (ii) we will refer to the three-dimensional quasi-monotone limiter (QMSL) as described in Bermejo and Staniforth [1992] as LQM3D (iii) we will refer to the split-dimensional version of LQM3D

as LQM and (iv) when LQM is applied in the horizontal directions only, without any limiter in the vertical, we will label this limiter as LQMH. The term split-dimensional LQM implies that the QMSL limiter is applied at each cubic interpolation separately i.e. every cubic interpolation in the stencil is limited before a further interpolation in the stencil is performed. Note that the term “limiter” or “monotone limiter” will mean the same concept.

4 Sensitivity tests from idealised cases

Here, “academic” case studies with “idealised dynamics” are presented which give useful information on the mass conservation and accuracy properties of the SL advection scheme and its various options. The advantage of these cases is that they present a simple controlled environment where we can expose the impact that different factors have on transport errors (advection in this case). Such factors are the terrain (flat versus orographic), the type of flow (convergent/divergent or divergence free), the direction of the flow and the location of the tracer.

4.1 Solid body rotation

The solid body rotation test described in Williamson et al. [1992] is a standard 2D (horizontal) test for checking the accuracy of an advection scheme as it has an analytic solution. It describes a divergence-free horizontal adiabatic flow in a spherical domain without orography (flat terrain) and without rotation. In its original form by Williamson et al. [1992], a tracer field in the form of a smoothly varying Gaussian bell, initially placed at the equator, was advected by a prescribed wind field performing one rotation around the Earth in 12 days. The test allows to specify the angle of the wind vector field with respect to the equatorial line changing the flow from pure zonal to pure meridional. For convenience and quick execution, the grid of the experiment was chosen to be tco79 (approximately 130 km horizontal spacing), but the conclusions hold for other resolutions as well. In this case we have excluded the COMAD interpolation method compared with the standard Lagrange interpolation method from our comparisons as there is no gain from it in divergent free flow cases.

The global error of the advection scheme and the cumulative global mass conservation error for a meridional advection case are plotted in Fig. 1. The term “advection global error” means that a geographically averaged mean absolute error norm is computed in the IFS Gaussian cubic octahedral grid i.e. the integral of the absolute value of the difference of the numerical solution (advected field) from its analytic solution. The cases displayed in Fig. 1 are: advection of a smoothly varying Gaussian bell and advection of a small plume tracer initial profile i.e. a disk with 6 degrees diameter with mixing ratio 1 inside it and 0 outside. Although this is a very simple case it reveals interesting facts:

- Both global error and mass conservation error are orders of magnitude larger in the plume tracer case. In contrast, errors in the Gaussian bell case are generally very small. For the plume tracer, errors remain low for wider radiuses (not shown here) but gradually increase as the disk’s radius decreases and collapses to a point. Inside a large-radius disk, the tracer field remains homogeneous, resulting in small errors, whereas at the front region (disk boundary), errors become significantly larger.
- The errors occurring in this simple case can be attributed to the semi-Lagrangian interpolation scheme errors. The slightly noisy error profile of Fig. 1c relates to different errors the interpolation method generates as the disk-shaped tracer moves

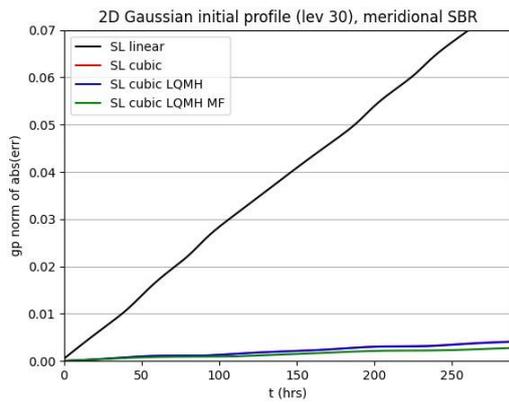
CATRINE

in a meridional direction of the quasi-uniform cubic grid (grid-points are not aligned in the north-south direction and the corresponding grid spacing changes slightly).

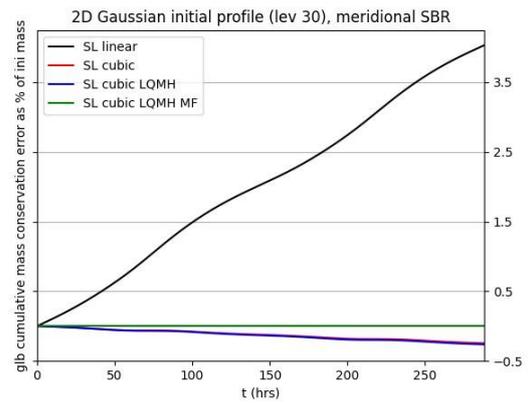
- Linear interpolation shows notably larger errors compared with cubic interpolation. It is highly diffusive and results in mass spreading (dispersion). The large diffusion effects of linear interpolation are clearly illustrated in the case of a single grid-point plume displayed in Fig. 2c,d. The impact is bigger in the meridional (slightly variable grid spacing with grid-points which are not aligned) rather than zonal case (uniform constant spacing grid). The apparent excellent conservation of the linear interpolation in the zonal flow case in Fig. 2b comes at the expense of heavily damped peak values and a very dispersive solution. So, while the conservation error has reduced the overall advection error has increased.
- The mass fixer has negligible impact in the accuracy but big impact in conservation i.e. mass conservation obtained without degrading accuracy as shown in Fig. 1a,c for both the Gaussian bell and the plume case.
- Not much sensitivity is found with respect to the limiter for the Gaussian bell but there is big sensitivity for the plume tracer, both errors are larger when the limiter is used. As the plume becomes smaller and collapses to a point the limiter produces more conservation error. The conservation error is much larger in the meridional direction than in the zonal direction as shown in Fig. 2a, b.

The solid body rotation case is a horizontal advection case. Cases where vertical advection dominates are presented in the next two subsections. This is worth exploring separately as there are differences in the vertical advection such as the use of time dependent levels (based on the surface pressure that changes in time), the stretching of the grid, the interaction with the boundary conditions and the existing large pressure (density) vertical gradient.

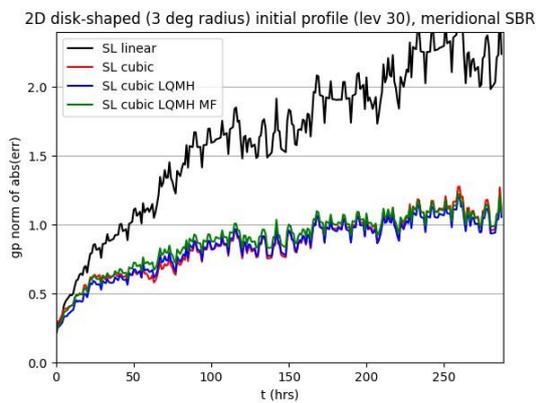
CATRINE



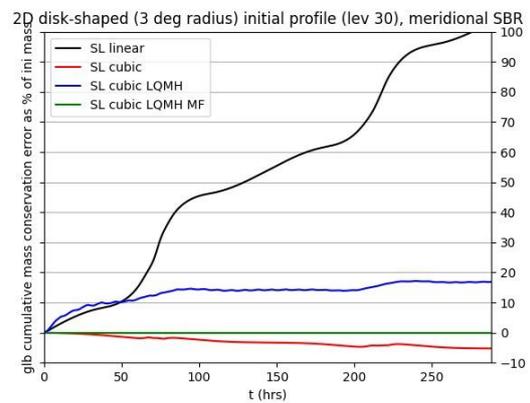
(a): Mean Abs Global error - Gaussian bell



(b): Mass cons error - Gaussian bell

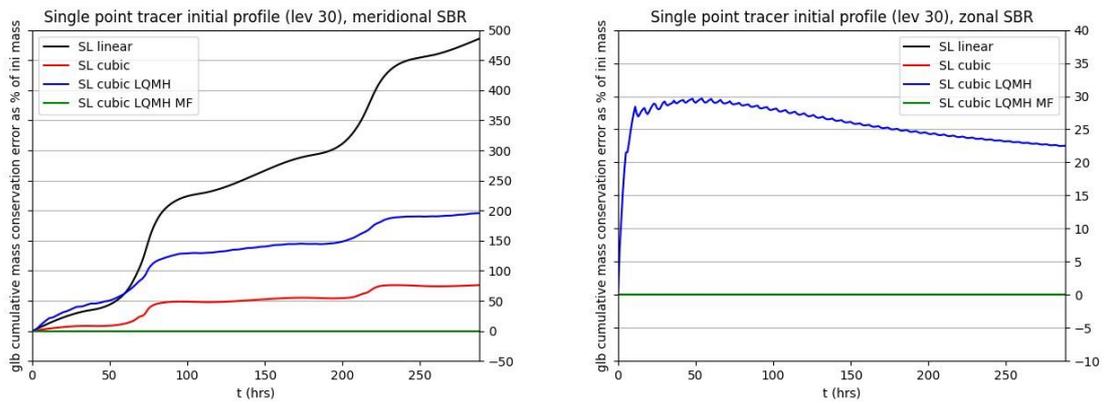


(c): Global error: 6 deg plume tracer



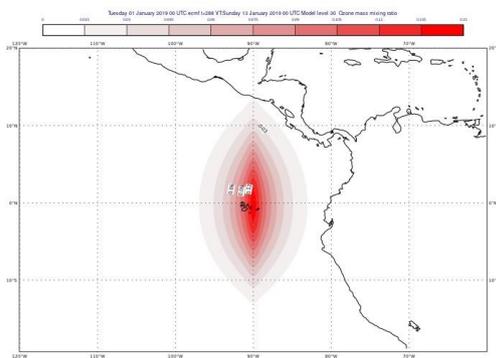
(d): Mass cons error: 6 deg plume tracer

Figure 1. Solid body rotation case in meridional direction. Left: Advection error i.e. absolute value of the difference of the advected field from the analytical solution after one rotation around the pole. Right: global mass conservation error.

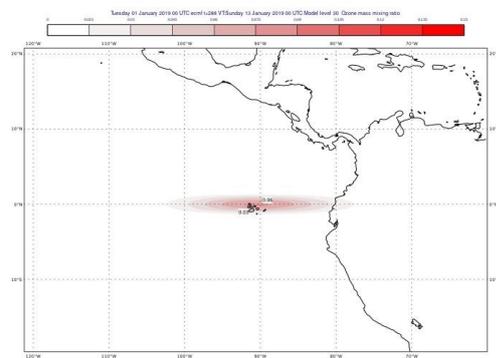


(a): Meridional flow

(b): Zonal flow along the equator



(c): Tracer after one rotation (meridional)



(d): Tracer after one rotation (zonal)

Figure 2. Advection of single point plume as in the solid body rotation case. Top: Global mass conservation error. Bottom: tracer at the end of a simulation (after one full Earth rotation) with linear interpolation scheme. Same scale is used in both plots, the maximum is limited from analytical solution value 1 to 0.15 to make it visible because of the large numerical diffusion of the linear interpolation. In (d) results from linear and cubic are not visible as they are both conserving mass. This is a favourable case for conservation as a single point tracer is advected in the regular East-West grid by a constant wind in this direction. Cubic interpolation produces undershoots/overshoots and it is corrected by the limiter which degrades conservation.

4.2 Single column rising bubble case

The idealised bubble case in Malardel and Ricard [2015], which has inspired the design of the COMAD interpolation (an option in the IFS SL advection scheme) presents a very specific and artificial situation of perfectly symmetric and converging horizontal winds toward a grid point at rest near the surface or near a rigid lid higher in the atmosphere (for example, the tropopause). In this configuration, the mass can grow infinitely when the model equations and their discretization scheme are not in flux form. Although this unbounded mass growth can only occur in such a “controlled” model environment, in practice aspects of this issue exist in real atmospheric flows, contributing to mass conservation errors. For example, consider a meteorological situation with light, converging winds at a nearly calm (no wind) grid point, which is at the model level closest to the surface. Consistently with mass conservation, this situation results in an upward motion above the calm grid point. In the SL scheme, the transport of tracer content

CATRINE

associated with the converging horizontal winds around the calm grid point is underestimated as the SL trajectory computation “sees” only the calm wind at the centre of the calm grid point. In the vertical, as the flow is ascendant above the converging zone, the upstream departure points is below the model computational domain (below the lowest atmospheric level) and the vertical interpolation degenerates to a simple copy of the original value at the level nearest the surface. The lack of horizontal mixing due to the underestimation of horizontal transport in grid scale (badly resolved) convergent zone, prevents tracer dilution, so the tracer mass continues to increase until the weather pattern changes. The COMAD interpolation addresses this very specific issue by modifying the interpolation coefficients so that they allow horizontal mixing with the neighbouring grid boxes in case of grid scale convergence.

The pathological behaviour of the IFS SL scheme near grid-scale convergence zone and the improvement when COMAD is activated are illustrated in Fig. 3 (the tracer bubble does not detach from the surface but fills the column). A warm bubble of $+3^{\circ}K$ perturbs an initial atmosphere originally at rest. The air of the warm bubble is filled up with a passive tracer of initial concentration equal to unity. Due to buoyancy, the bubble rises, and the tracer is entrained with warmer air. As buoyancy effects become dominant for a small Rossby radius of deformation, the “small planet” configuration with a planetary radius divided by 15 is used to increase resolution without increasing the computation time. The simulations are performed without rotation with the adiabatic configuration of the IFS using a linear or a cubic grid with a T511 spectral truncation (i.e. a horizontal resolution ~ 2.5 km for the linear grid, ~ 1.3 km for the cubic grid), 90 vertical levels, and a time-step of 60s. The simulated time is 6 hrs. The differences between linear and cubic grid are explained in [Malardel et al., 2016].

Fig. 3 shows vertical cross-sections of tracer concentration for the original SL method and for COMADH (COMAD for horizontal interpolations only), both for linear and cubic interpolation options with the LQM limiter. As expected, the linear interpolation (order 1) is more diffusive than cubic interpolation (order 3), whether COMAD is used or not. The tracer concentration is consequently lower in the wake of the bubble when COMAD is used, especially in the case of linear interpolation. The gain is larger for linear interpolation for which, conceptually, the COMAD correction mimics a flux-form, finite-volume scheme along both horizontal direction of interpolation.

Results in Fig. 4 show that COMAD improves conservation for both a linear grid (smallest convergence feature of size $1\Delta x$) and a cubic grid (smallest convergence feature of size $2\Delta x$). So, even if COMAD was designed to address a limitation of the SL at the grid scale, it still improves conservation in the bubble case with the cubic grids currently used in the IFS [Malardel et al., 2016] for which the spectral divergence field is filtered at $4\Delta x$. Notably, with or without COMAD, the mass conservation error is smaller on the cubic grid compared to the linear grid, with a reduction by a factor of 3. When cubic interpolation is used, combining COMAD scheme with the cubic grid, yields a clearly favourable result reducing the overall mass conservation error by a factor of 7 compared to the case without COMAD on the linear grid. The cubic grid is thus the preferred choice for weather and atmospheric composition forecasts. It is also important to note that these errors are intentionally exaggerated, as the test case is designed to produce this effect. Here, as in other cases presented in the following sections, mass growth occurs gradually. While the mass conservation error on the cubic grid remains less than 1% at each timestep, the cumulative effect leads to an exponential increase in error when measured relative to the initial tracer mass. In realistic simulations, errors are much smaller, and furthermore the mass fixer is activated which prevents such spurious growth.

CATRINE

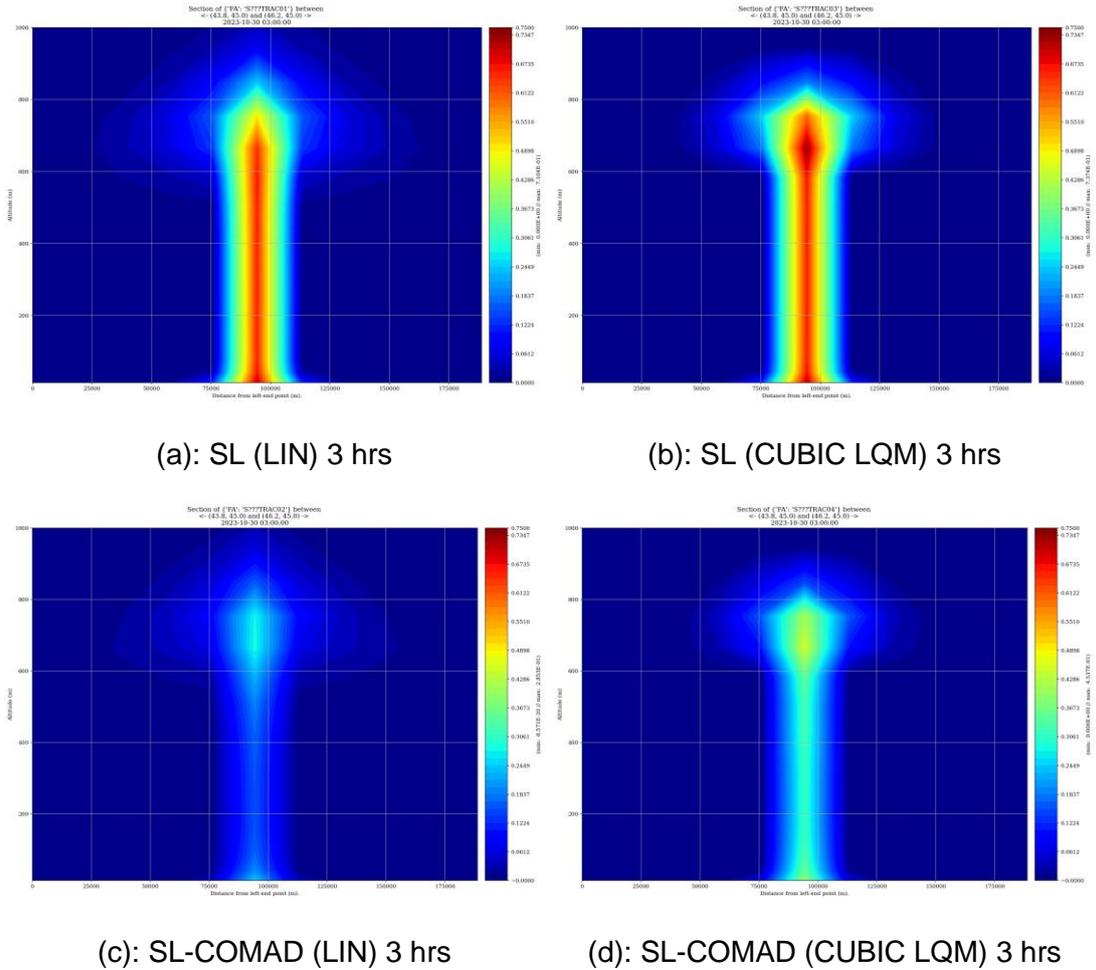


Figure 3. Vertical cross section of tracer concentration after 3 hours of simulation (cubic grid T511 truncation). (a),(b): original SL method. (c),(d): SL-COMAD method using linear interpolation ((a), (c)) or cubic interpolation with LQM limiter ((b),(d)).

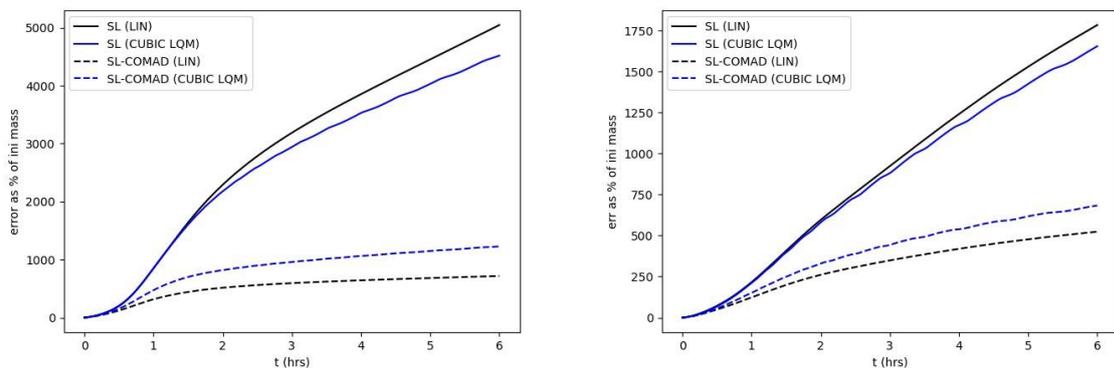


Figure 4. Mass conservation error growth (in percent of initial tracer mass) for the bubble test case. (Left: linear grid. Right: cubic grid).

4.3 Advection over a mountain range

Idealised test cases which have originally been designed to study mountain waves at very high resolution in mesoscale regional models have been adapted to investigate the role of the bottom boundary and the deformation of the terrain following levels on the transport conservation error at resolutions currently used by global NWP models.

For these series of test cases, mountains of different shapes and heights are centred at the equator. An initial non-sheared zonal flow (10 m/s at the equator) in solid body rotation is in geostrophic equilibrium with the initial pressure field. The temperature profile is isothermal. A disk of tracer of 100 km radius (specific ratio = 1 inside the disk, 0 outside) is initially positioned about 500 km upstream of the orography.

The orography is setup as an elliptic mountain of 100 km width (East-West direction) and 250 km length (north-south direction). The horizontal spectral resolution of the simulation is TL639 (same as ERA5 resolution, i.e. about 30 km horizontal mesh resolution) and 105 hybrid terrain following pressure levels (η -levels). The lowest 84 levels follow the orography as, by construction, the hydrostatic pressure of these levels has a contribution from the surface pressure. Above, the levels are pure isobaric levels. The contribution of the surface pressure in the first 84 levels is the strongest near the surface and slowly weakens with height. The vertical advection is computed with respect to the hybrid η -levels using the corresponding vertical velocity which depends on the horizontal divergence of the winds, the surface pressure and its tendency, the dot product of the horizontal wind with the horizontal surface pressure gradient as shown by equations (2.18) and (3.6) in ECMWF, 2024 (IFS documentation). This equation reveals that the terrain has large influence in the vertical velocity, particularly in the levels closer to the surface. This influence completely diminishes in the stratosphere as the levels become independent of the terrain.

With this test case, we studied the sensitivity of the tracer advection, in particular its dispersion, with respect to the initial elevation of the tracer and to the maximum height of the mountain and then we focus on mass conservation.

Three different elevations have been tested:

- the initial disk of tracer is placed at the first model level above the surface (SURF)
- it is placed at about 4000 m in the troposphere (MID level, non-isobaric)
- it is placed at the first isobaric level (TOP level)

Even with a 1 km high mountain, the flow at MID and TOP remains almost horizontal. As TOP is an isobaric “quasi-horizontal” level, most of the tracer stay at TOP during the simulation.

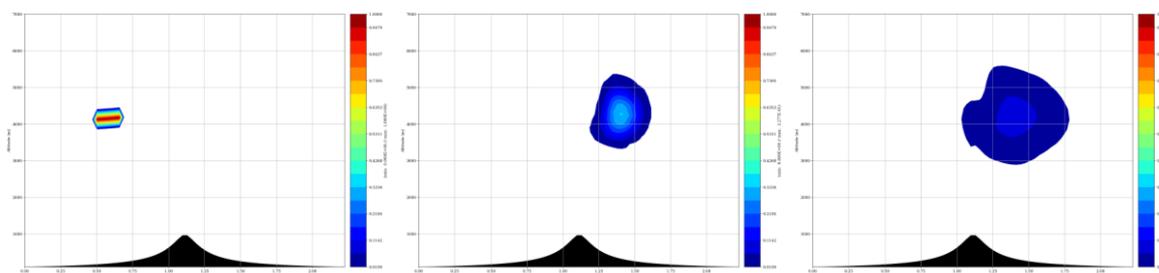


Figure 5. Vertical cross-section along the equator for a tracer initially at MID level (case with a 1km high mountain). Horizontal axis: horizontal distance from left point. Vertical axis: altitude (m). Initial state (left) and after 24 h of simulation using the cubic interpolation+LQM limiter (middle) or the linear interpolation (right).

CATRINE

At MID level, the horizontal transport of the tracer means that it must be first advected downward with respect to the terrain following levels before the ridge and then upward after the ridge (around MID, the levels follow the orography but the tracer goes horizontally, a tracer moving horizontally is then advected towards lower η -levels when the slope of the level is up and then it reverses when the slope of the level is down). Because of this vertical advection, there is more vertical diffusion towards the neighbouring levels than there is at the TOP case, in particular if the linear interpolation method is used (Fig. 5). This is a source of error and potentially of conservation error.

At SURF, the tracer “glides” along the mountain. It mostly remains at SURF (even if the air goes up, there is almost no vertical advection in terms of computational space vertical velocity η as the model levels around SURF follow the orography). In the 1 km high mountain case, the tracer is partially blocked and then flows around the mountain. For the 500 m and 100 m high mountains, in agreement with the Froude number of this case, it mostly goes above the ridge. This is shown in Fig. 6 where the strongly diffusive nature of linear interpolation is also evident.

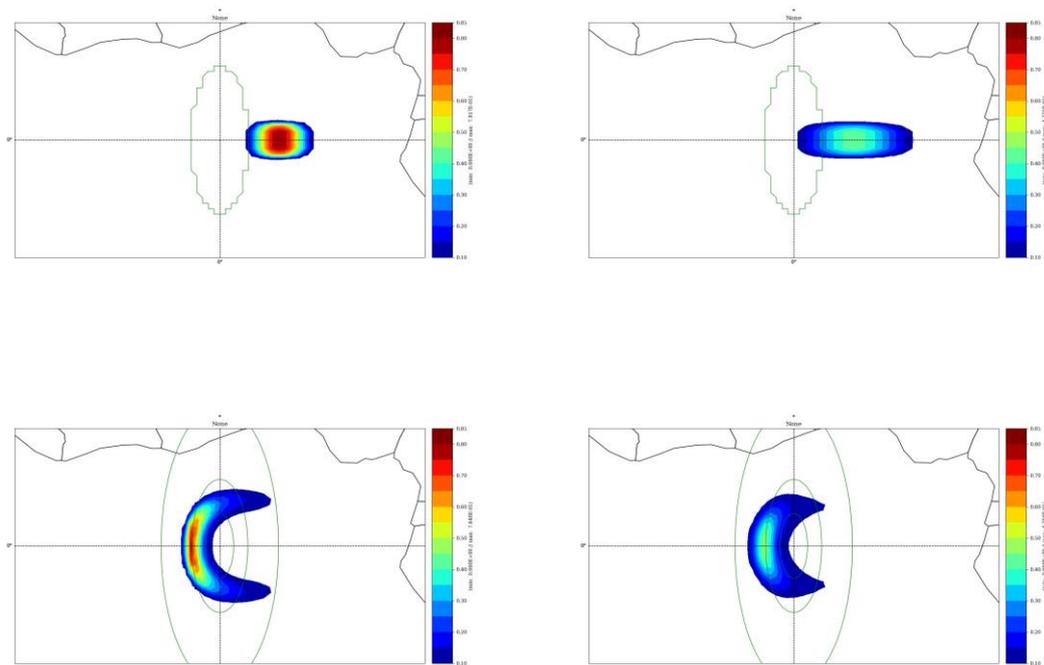


Figure 6. Map of tracer concentration at SURF after 24 hours of simulation for (top) a 100m, (bottom) a 1 km high mountain (green lines gives de mountain height from level 50m, every 250m). Left: cubic interpolation with LQM limiter. Right: linear interpolation.

Tracer mass conservation errors are very dependent on the initial level of tracer, on the type of interpolation selected in the SL scheme as well as on the height of the mountain. From Fig. 7, we conclude that:

1. Far above the surface (TOP, isobaric level), the conservation is very good. The deformation of the terrain following level is then a source of error in a horizontal flow. Errors (in percent of the initial mass) are worse when the tracer is at SURF initially.

CATRINE

2. The cubic interpolation without a monotone limiter is the most conservative but there is a creation of unphysical negative concentrations. Because of this and a tendency of the cubic-interpolation to undershoot in such “plume-type” tracers, near the surface, unlike all other options, there is a small mass loss in this case.
3. The results with cubic interpolation and the monotone limiter LQM show errors which are related to the vertical advection across hybrid η – levels upstream and downstream of the orography in the MID case. In the SURF case, errors grow only when the tracer is downstream of the main ridge.
4. With the linear interpolation, there is a large error growth for SURF, even if this interpolation method doesn’t need a monotone limiter. This is something which is also observed in more realistic test cases, and it is related with the terrain deformation (see item 1) as well as the numerical treatment of the vertical boundary condition (see section 5.3). Furthermore, near the surface solid boundary, the wind convergence/divergence is higher than in regions away from it and as we saw in the bubble case this situation can contribute to conservation error. For MID and SURF, the conservation errors with linear interpolation are of the same order of magnitude than with cubic interpolation and limiter on the lee side of the mountain. They are smaller upstream.

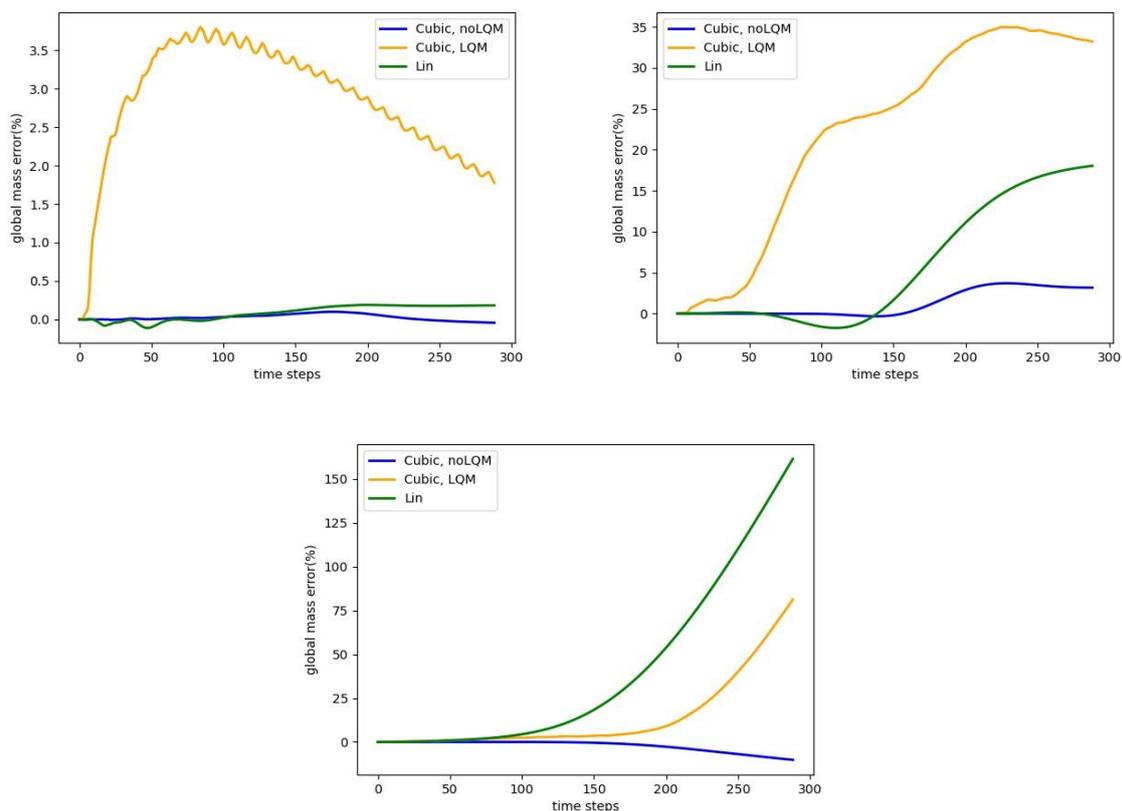


Figure 7. Evolution of the total mass of tracers for advection over a 1km mountain. Clockwise: TOP, MID and SURF 24hrs simulations (x-axis in time steps). The y-axis gives the percentage of error with respect to the initial mass given by equation (5). Blue lines: cubic interpolation without monotone limiter. Orange lines: cubic interpolation with LQM limiter. Green lines: linear interpolation.

CATRINE

The impact of the height of the topography in term of conservation error is shown on Fig. 8 for the SURF case. The higher the topography, the larger the error for all interpolation choices (cubic without limiter, cubic with LQM limiter and linear). For flat terrain, the conservation error is very small. Cubic interpolation generates negative errors which are somehow compensated by the monotone limiter. The activation of the current version of COMAD (see section 4.2) to correct SL mass error in convergent flow has very little impact in these cases (not shown).

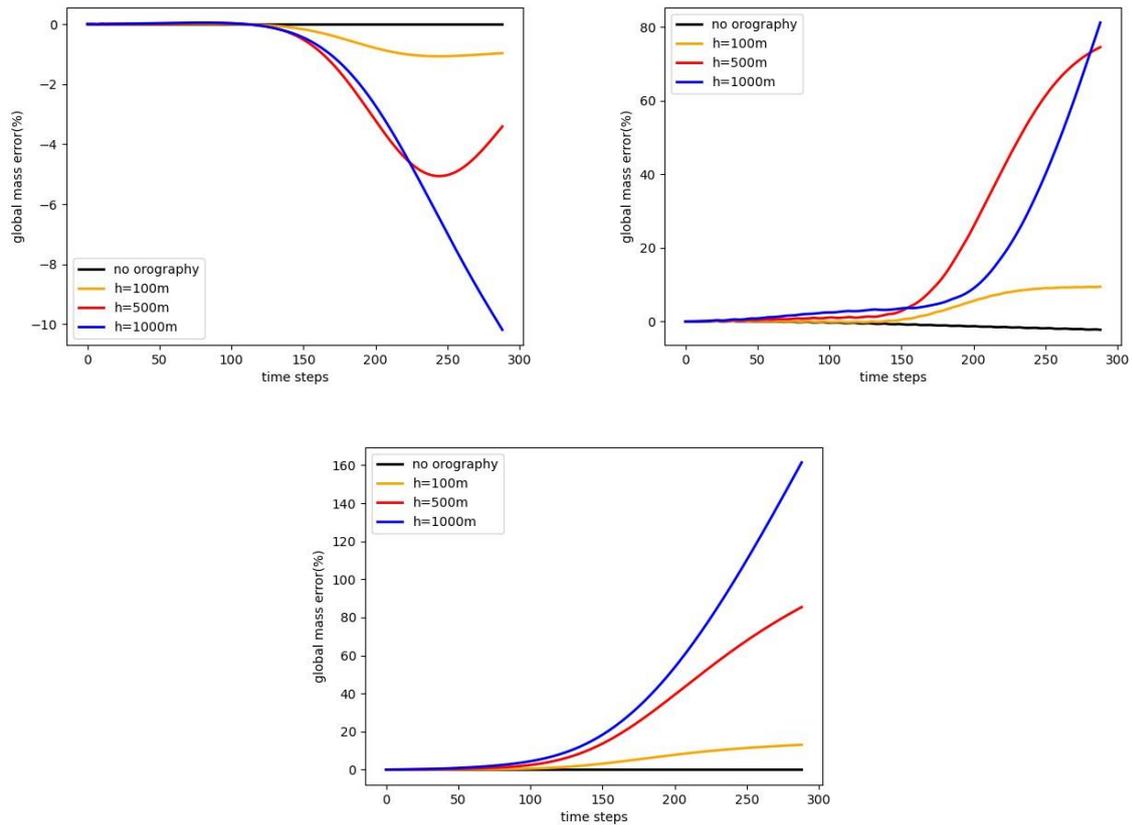


Figure 8. Evolution of the total mass of tracers initially positioned near the surface (SURF) for (clockwise): (a) cubic interpolation without monotone limiter, (b) cubic interpolation with LQM limiter and (c) linear interpolation during the 24 h of simulation (x-axis in time steps). The y-axis gives the percentage of error with respect to the initial mass. Blue lines are for a 1 km high mountain, red for 500 m, orange for 100 m and black for flat terrain.

5 Advection sensitivity tests with plume type tracers in real weather scenarios

While academic tests are informative, they don't fully represent the flow complexity and specific tracer characteristics (plume types emitted from a point near the surface) which are important for CO2MVS. To thoroughly evaluate IFS's behaviour, we developed a tracer test bed that assesses passive advection of multiple artificial tracers — isolating dynamics from physical parameterizations — using winds from a real forecast setup. It was initially inspired by the setup of Eastham and Jacob [2017] but has been extended to offer few extra capabilities such as passive advection of plume-like tracers which can be initialized at any region of the atmosphere and advection of a passive tracer emitted from a single or multiple user-defined source points in the globe. It is worth emphasizing that the cases presented here are designed to stress test the SL advection scheme and its components such as mass fixers and monotone limiters. The mass conservation errors produced by the cases we examine in this section are much larger than corresponding errors of realistic scenario greenhouse tracer simulations presented in Section 6. The main reasons are that in realistic simulations: (i) physical processes such as the parametrized turbulent mixing smooth tracer gradients (ii) the strength of emissions is smaller than the ones considered in artificial tests here and (iii) the background tracer value is very large while there is none here.

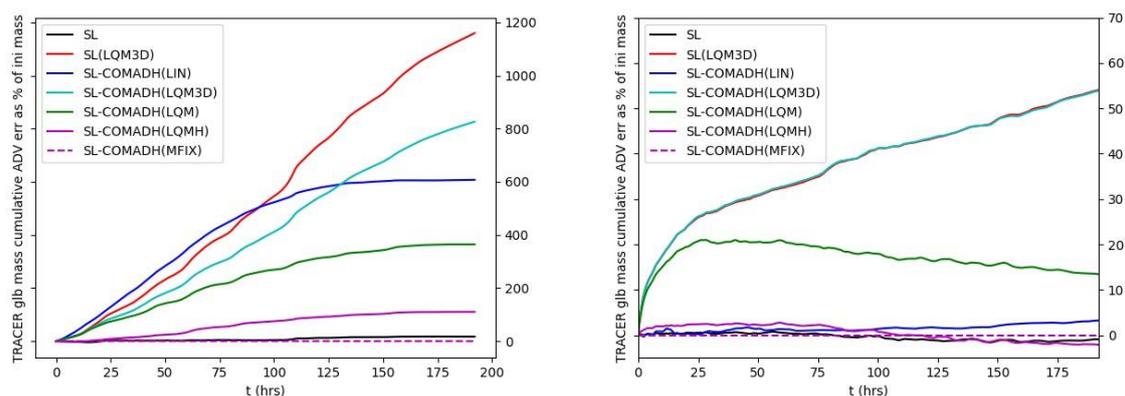


Figure 9. Mass conservation error growth (percent of initial tracer mass) for a 4x5 degrees tracer transported by IFS model winds over 8 days. Left: tracer initialized near the east coast of China at the near-surface level. Right: tracer at the same location but in the stratosphere near 30 hPa. MFIX is run with a mass fixer. Comparisons are made between standard SL advection cubic interpolation scheme (SL) and COMAD linear (LIN) and cubic with different limiters.

5.1 Large plume tracers without emissions

The first case study examined is similar to the case described by Eastham and Jacob [2017]. A rectangular tracer of 4x5 degrees in longitude/latitude, resembling a plume, is initialized at different heights (model vertical levels) at different parts of the globe. More specifically, we have experimented with locations in the mid-latitude zones of strong industrial emissions and the tropics and compared results when the tracer is initialized at the near surface atmospheric level, at a model level near the 850 hPa level (near cloud top), at the top of the troposphere near 250 hPa level and in the stratosphere at 30 hPa. In the interior of the rectangular tracer, the concentration (mixing ratio) increases abruptly to a fixed value while outside is 0. As expected, and experimentation confirms, with a zero-background value, the same behaviour in mass conservation error growth is obtained regardless of the precise value of the concentration inside the plume. Hence for

CATRINE

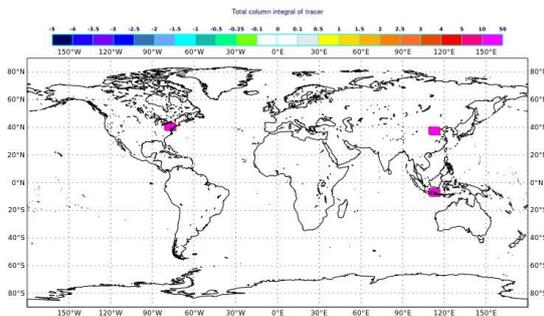
simplicity, given the idealised nature of experiments, we have set the mixing ratio to 1. In this case, the tracer is like a step function, which is challenging for any known SL interpolation with order higher than one (linear). High-order methods are known to undershoot or overshoot in such situations, generating unphysical values such as negative concentrations or new minima and maxima in the field. The interpolation limiter prevents this from happening, but it has an impact on conservation. As we will show, it significantly amplifies the growth of mass conservation error, as applying it sacrifices any conservative property of the pure Lagrange polynomial interpolation. However, without the limiter, large undershoots or negative mass can occur, artificially offsetting positive mass growth.

An example of mass growth due to lack of conservation in SL advection is shown in Fig. 9. In this case a tracer initialized in Eastern China (see Fig. 10a for the precise location) is advected by winds from a forecast with 25km horizontal grid spacing and 137 vertical levels. The initial date for this forecast is 1/07/2022. No source terms and physical processes act on the tracer so any mass growth observed is entirely due to the model dynamics. The methods tested are summarized in Fig. 9. A striking feature of this plot is that when a tracer with identical shape and extend is initialized at the near-surface level the mass conservation growth is at least one order of magnitude larger than when it is initialized in the stratosphere (30 hPa level). For the near-surface initial tracer profile, the conservation error in the simulation without any limiter (black line) is very small, more than two orders of magnitude smaller than when the 3D limiter is used. The same is true for the equivalent COMAD run without limiter (not shown here). Different types of limiters produce different rates of growth with LQM3D being the worse while LQM is reducing the conservation error by a factor of 3 compared to LQM3D. When the limiter is applied only in the horizontal (LQMH, purple line) then the mass conservation error reduces substantially, therefore, filtering of vertical advection has a big impact in producing conservation error. COMAD seems to reduce the mass conservation error compared to standard SL interpolation scheme, however, this is still very large. These conclusions are robust as they can also be verified in other runs where tracer is initialised at different surface locations. It is worth mentioning that the chosen case in China presented here is one of the worst cases in terms of mass conservation error growth. The limiters have similar effects in the stratosphere (see Fig. 9b) although the mass conservation errors generated are much smaller. The simulation becomes nearly conserving without limiter or when it is applied in the horizontal only. The linear interpolation produces very small conservation errors. The stratospheric case resembles the conditions of the solid body plume test discussed in section 4.1 and some of the mass conservation profiles obtained here are similar with those in Fig. 2d. Furthermore, the smaller conservation errors observed in the stratospheric case, agree with the mountain case analysis of the previous section where we saw that in the horizontal stratospheric levels the conservation errors are small.

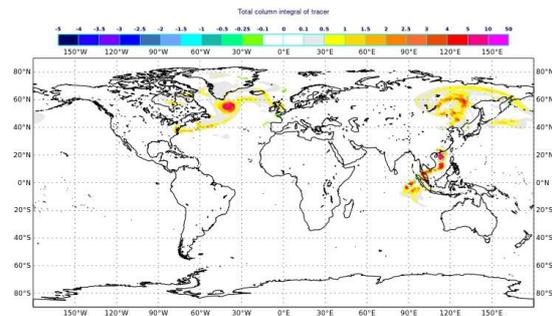
A global map which visualizes the impact of mass growth due to lack of conservation can be seen in Fig. 10. The evolution of the total column mass in kg/m^2 for three plume tracers is illustrated at 7 days after the initial forecast time. The impact of different limiters applied in COMAD cubic-Lagrange interpolation is examined there. It is evident that although the tracer has spread in the same area, forecasts with limiter produce much higher total column values, especially the one with LQM3D. In contrast, without limiter, significant undershoots are produced during advection which result to negative (unphysical) concentration values. With the mass fixer active, any sensitivity in total column mass with respect to limiters reduces considerably and results are similar to these without limiter but without having unphysical negative values. So, the mass fixer seems to restore conservation without causing a big change in the solution of the transport equations, it looks like it acts where it should. A time series of the global mass conservation error corrected by the mass fixer at each time-step, for two tracers (one initially in China and one in Indonesia) is shown in Fig. 11. We notice that initially, the magnitude of the

CATRINE

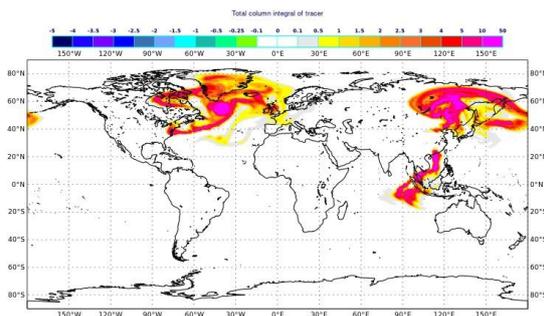
correction is of the order of 1% of the global mass of the tracer but gradually, as the tracer spreads in the atmosphere and its gradients weaken it reduces to 0.1% or less. An interesting diurnal variation is also shown in the case in China: it is much smaller during the day than the night where winds are near calm. In general, placing the tracer in a zone of active horizontal winds, even at the surface we observe reduction of the conservation errors. The same analysis was made in the idealised bubble case.



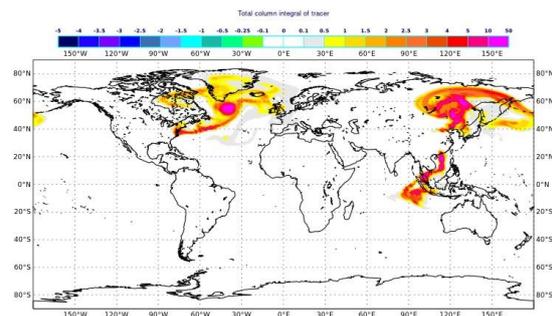
(a): t=0 (Initial Values)



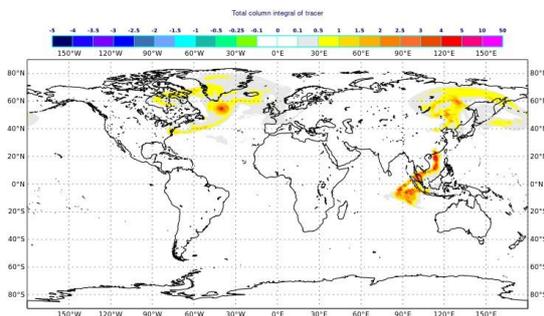
(b): t=168 hrs with COMAD (no limiter)



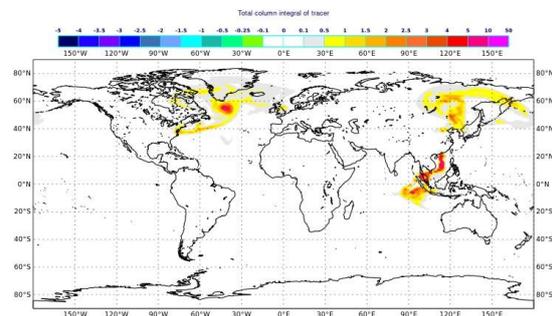
(c): t=168 hrs with COMAD (LQM3D)



(d): t=168 hrs with COMAD (LQM)



(e): t=168 hrs Mass Fix + COMAD (LQM3D)



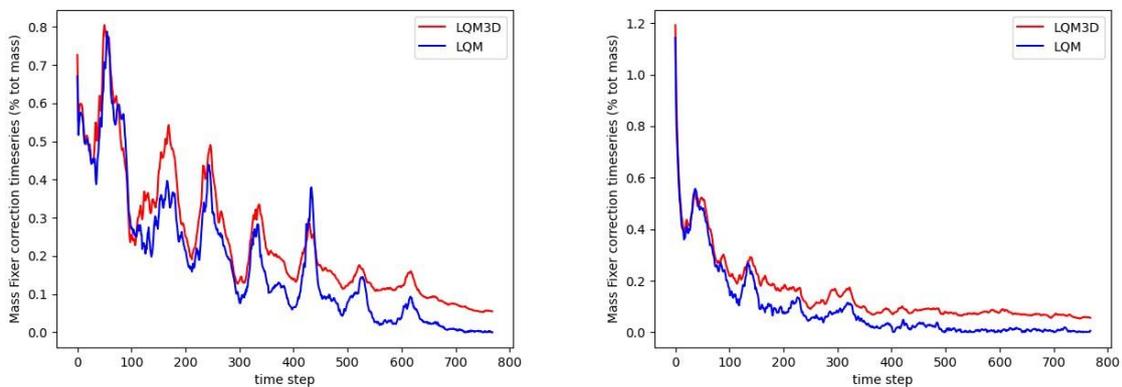
(f): t=168 hrs Mass Fix + COMAD (LQM)

Figure 10. Total column integral (kg/m^2) of rectangular shape plume tracers defined initially at a single model level and advected with COMAD option in IFS and with different limiters as indicated in parenthesis and without/with Mass Fixer. There are three tracers in the simulation: one in eastern China, one in Indonesia and one in eastern USA. All

CATRINE

tracers all initialized at the near surface atmospheric level (137) which has approximately 10m height.

The persistent undershooting which is exhibited during the advection of such type of plume tracers results in excessive mass growth: when an undershoot occurs when interpolating the tracer field at its departure points, the limiter locally clips the interpolated value to the local minimum. This clipping effectively increases the concentration and therefore it is equivalent to adding mass locally. This shows clearly in Fig. 12 where the amount of tracer the limiter clips has been accumulated in time and the total column equivalent is plotted. It is interesting to observe that for plumes resembling a step function from 0 to 1, as in Fig. 12a, persistent undershoots occur, while for plumes resembling a step function from 1 to 0 i.e. a tracer with initial concentration 1 everywhere except for a small rectangular area, which is the case in Fig. 12b, persistent overshoots occur. The former, which is a typical setup, results in increased mass (positive conservation error) while the latter would have resulted in reduced mass (negative conservation error).



(a): Tracer initialized in China (surface). (b): Tracer initialized in Indonesia (surface).

Figure 11. Time series of global mass conservation error calculated and corrected by the mass fixer at each time-step for the two forecasts displayed in Fig. 10 (e), (f). The error is expressed as a percent of the global mass of the tracer at each time-step. Red lines: COMAD with LQM3D limiter. Blue lines: COMAD with LQM limiter.

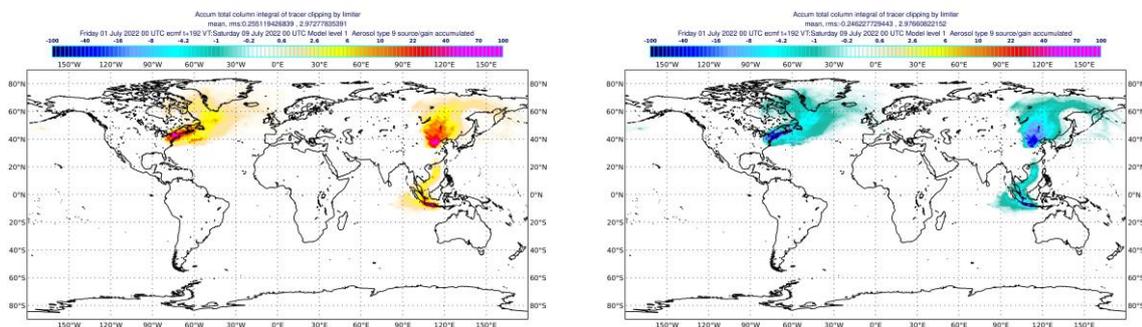


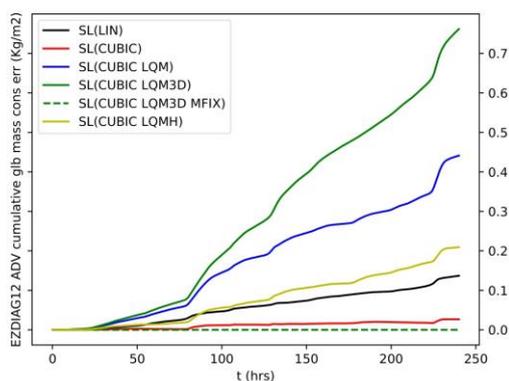
Figure 12. Total column integral (kg/m^2) of tracer concentration clipped by the LQM3D limiter and accumulated in time for three tracers in Fig. 10. Left: A rectangular tracer with initial concentration 1. Right: a tracer with initial concentration 0 in the same rectangular area and 1 elsewhere.

5.2 Plume tracers emitted by a single-point source

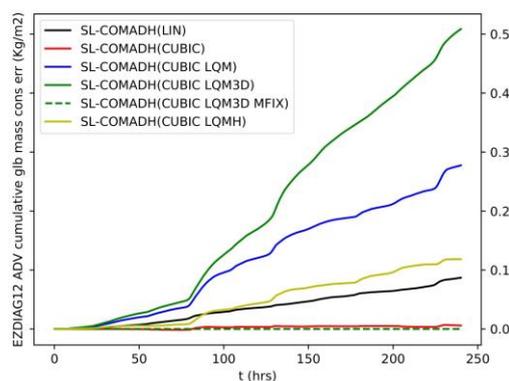
The case study presented in Section 5.1 represents passive advection of a plume-type tracer which consists of a group of adjacent grid points, for example, it could be regarded as pollution from a large city with an extended industrial area transported by the winds. To study the IFS behaviour, while trying to emulate in a simplified way the important to CO2MVS case of point source emissions, we have constructed a case with 18 tracers emitted from different areas (tropics/midlatitudes, land/sea). Some of the emission points are located near known industrial sites. Each tracer has initial concentration 0 which is increased by a fixed amount of 0.5 kg/kg at each time-step. This new setup represents a very strong emission scenario. Considering that the tracer is emitted over 900s in a grid-box approximately 25km x 25km, this is comparable with the strength of a strong volcanic emission such as Mt St Helens eruption (10^7 kg/s) rather than of an anthropogenic emission, and it is challenging in terms of conservation for the SL advection scheme as the tracer field starts as a single-point discontinuity in the atmosphere. The emitted plume is advected by the winds without any physical process acting on it, making it even more challenging, as in practice parametrized turbulent mixing would smooth tracer concentration gradients and would reduce mass conservation error growth. Essentially it is another case of stress testing IFS advection under the most unfavourable conditions for conservation.

This case is run for two separate dates (summer/winter): 01-07-2022 and 01-01-2023. Various sensitivity tests have been conducted on this case to examine the sensitivity of advection scheme options such as interpolation method order (linear/cubic), the interpolation method (standard Lagrange or COMAD) and its limiter. The accumulation of advection mass conservation error in kg/m² (i.e. vertically integrated) of a tracer emitted from a location in China is summarized in Fig. 13. When the mass fixer is activated (MFIK dashed green line) then the accumulated mass conservation error is equal to 0. However, this is not the case without mass fixer. The behaviour with respect to limiters and COMAD is similar to the behaviour observed in the case of rectangular plume analysed earlier in this section. COMAD performs clearly better but nevertheless the mass growth is still large. The limiters have a big impact in mass conservation with LQM3D being the least conserving. LQM reduces by approximately 50% the error. The simulations with horizontal limiter LQMH (no vertical) are more conserving (however does not filter all negative concentrations). Linear interpolation shows clearly smaller errors.

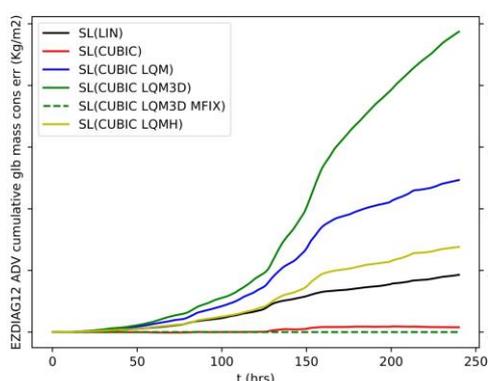
CATRINE



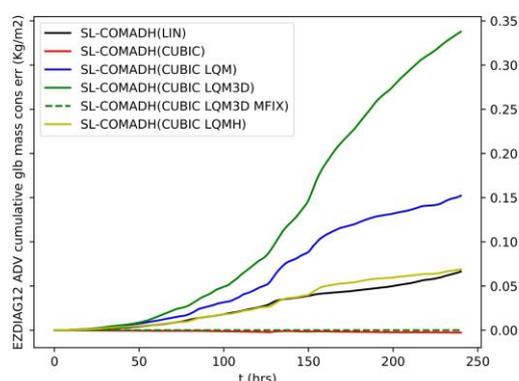
(a): Winter case - SL options



(b): Winter case - SL-COMAD options



(c): Summer case - SL options



(d): Summer case - SL-COMAD options

Figure 13. Accumulation of mass conservation error in kg/m^2 in a case where a tracer is emitted from a point in China with coordinates (113E,37.5N). The emission occurs at the near-surface atmospheric level of the IFS (level 137). Same options of the SL advection scheme as in Fig. 9 are tested and compared here.

In Fig. 14, we have aggregated the results from the 18-point source tests for the winter and summer simulations and for different options in the advection scheme (using linear interpolation or cubic with LQM/LQM3D limiter). The linear interpolation scheme exhibits the lowest conservation error, and from the cubic interpolation scheme family where a limiter must be used to enforce conservation, the LQM limiter produces smaller conservation error. However, as it was previously shown, despite the good “conservative” performance of the linear scheme on this point source emission test, overall, it is not a good option as it produces very diffusive and less accurate results which is also confirmed in this case (not shown here). In agreement with previous sections, we find that the largest conservation errors are produced both in calm wind situations at flat terrain as well as in situations of flow over steep orography in both hemispheres and in the tropics.

CATRINE



(a): SL-COMAD scheme summer summary

(b): SL-COMAD scheme winter summary

Figure 14: Comparison of mass conservation error performance of point source emission tests with respect to interpolation order (linear vs cubic) and limiters (LQM vs LQM3d) and across different locations in the globe. In each bar, results from a specific interpolation method/limiter are gathered showing the frequency where each method was the most conserving.

5.3 The impact of the lower boundary

The bubble case study in Section 4.2, has demonstrated that in a transport scheme as IFS one which is not in flux form, the interaction with the boundary can be an important source of mass conservation error in situations with near calm winds. It is worth exploring if this aspect is important in situations of real atmospheric flows and how much it contributes to mass conservation error. To answer this question, we have repeated the point source experiments described in Section 5.2 but changing the vertical location of the emission point. It is worth reminding here, that as mentioned in Section 4.2, a special interpolation procedure is applied when the upstream departure points cross the lowest vertical level (first model level above surface). In this case the tracer is not interpolated (no field value known below the lowest level), neither extrapolated (the option exists but makes the conservation worse) at the departure point location below the lowest level but the lowest level is used. However, when this situation occurs persistently at a location, it can increase the mass of a tracer (e.g. bubble case) unless a proper flux formulation is used which isn't the case in SL advection.

In Fig.15 we compare the mass conservation error for the case of Fig. 13d when the emission point is set at level 137 (near surface) and when it is set 8 levels above the surface (level 129). This is a near calm wind meteorological situation. It is evident from these plots that when the emission is near the surface the mass grows by a factor 2 to 3 more than when it is far from the surface. This increase is consistent for both linear and cubic interpolation with both limiters and it can be observed in all emission points. Furthermore, we have repeated the test where the entire atmospheric level above the surface is filled with tracer and compared it with the case where the tracer is placed at a higher level above noticing a significant reduction in the conservation error. This sensitivity exists either with standard Lagrange interpolation or COMAD, but as demonstrated in the bubble experiments, COMAD reduces this problem. A note of caution is that in the test presented here, that the advection of two tracers with identical mixing ratio at the same geographical location but at different levels is not directly comparable given the different ambient conditions (wind and air density). Therefore, changes in mass conservation error

cannot be entirely attributed to boundary effects. However, the higher mass growth for tracers near the lower boundary is a systematic error which is not easy to manage in a SL scheme (no flux form and neither finite-volume). Ideally, we would like to improve COMAD further to be able to handle this effect better.

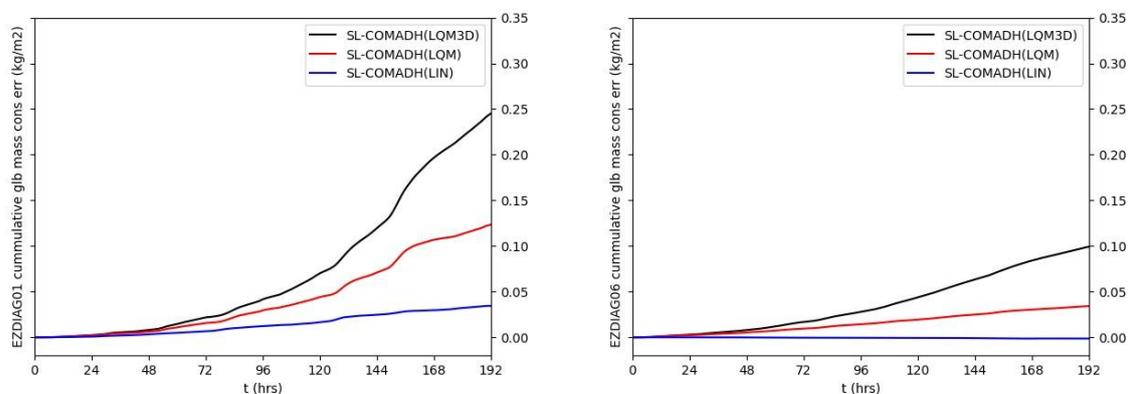


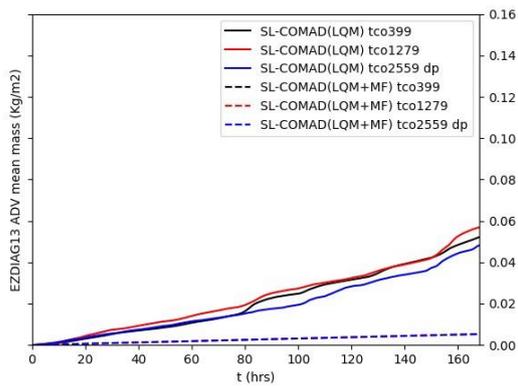
Figure 15: Artificial accumulated mass growth in kg/m^2 for the summer case displayed in Fig. 13 and for COMAD scheme. Left: emission from near surface IFS model level (137). Right: emission from a point 8 model levels above (129). Note the reduction in mass growth in the right plot.

5.4 The impact of resolution in mass conservation error

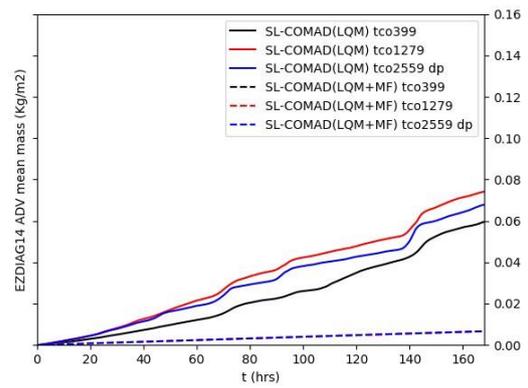
Another question that is worth exploring is whether increasing horizontal resolution significantly increases the mass conservation errors and has an impact on simulations with the mass fixer. Past studies in 9km and 25km resolution (see Agusti-Panareda et al. [2017]), have shown that the fixer performs well in higher resolution, improving the quality of the results. To explore this question further, we have repeated point source emission tests at different resolutions: 25km, 9km, 4.4km. At each time-step the concentration at the emission grid point is increased by a constant to make sure that the same amount of mass is injected in the atmosphere. This constant increase factor is normalized considering time-step and the volume of the grid-box. The 18 different locations around the globe which have been previously tested have been used.

Representative results from 4 different point sources (Indonesia, the northeastern US, the Baltic Sea, and the northwestern Australia) have been selectively plotted in Fig. 16 where the mean tracer mass (in kg/m^2 units) is plotted (rather than the accumulated mass conservation error plotted previously). All emissions are placed at the near-surface IFS model level (137). The forecasts are initialized from the ECMWF operational analysis dated 01-07-2022 and run for 7 days. With the mass fixer, all simulations show the same mass growth, which is attributed to the emissions inject into the atmosphere. Without mass fixer: apart from 3 out of the 18 cases, no big differences in mass growth among different resolutions have been observed. Regarding the three outlier cases, in two of them the mass conservation error has increased with resolution, in one it has reduced. The illustration in Fig. 16 (d) is the most noticeable outlier among the cases examined and shows a large difference (nearly by a factor of 3) in mass growth between 25km and 4.4km. However, even there, when the mass fixer is activated, the total mass of tracer increases as expected from the injection rate. Finally, it is worth mentioning that no difference has been seen in the results between single and double precision runs.

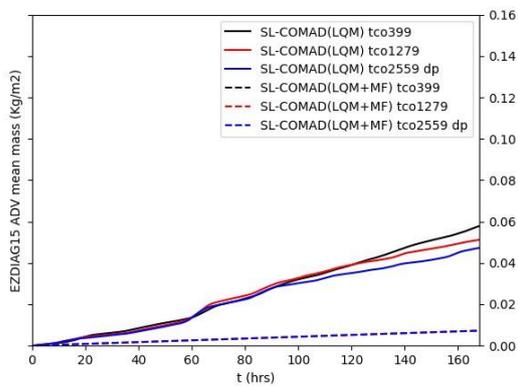
CATRINE



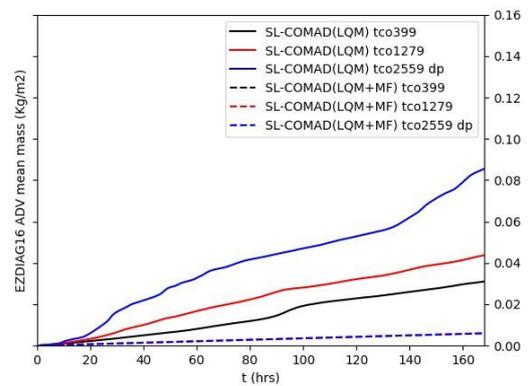
(a): Emission point in Indonesia



(b): Emission point in northeastern US



(c): Baltic Sea emission point

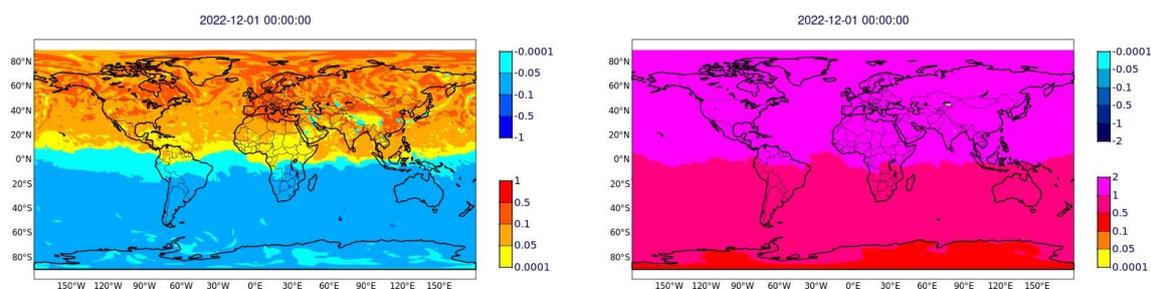


(d): Emission point in northwestern Australia

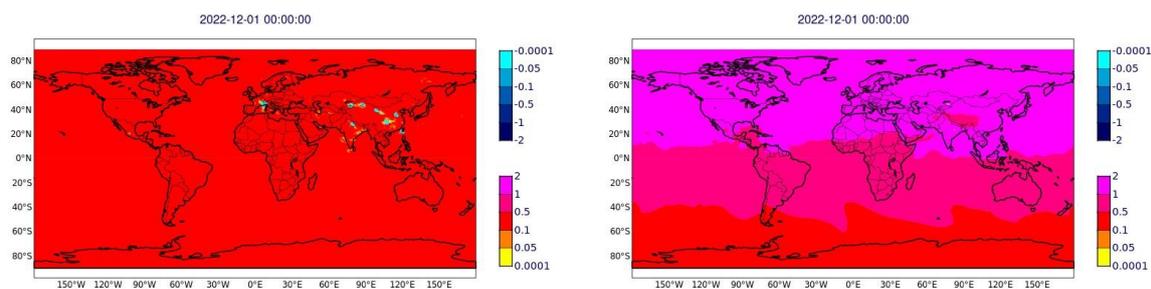
Figure 16: Mean mass (kg/m^2) from tracer emission simulations at different horizontal resolutions (grid spacing): 25km (black solid line), 9km (red solid), 4.4km (blue solid) grid spacing. The dotted lines represent mean mass from equivalent runs where the mass fixer has been activated. Here 4 different tracers are compared at different resolutions each emitted from a different point. All runs with COMAD interpolation and LQM limiter.

6 Advection sensitivity tests on CATRINE simulations using inter-comparison protocols

Global model emission protocols for model inter-comparison studies have been defined in CATRINE WP5 and WP7. The tests and the analysis of results conducted in previous sections have shown that the limiter and the mass fixer are two key components of the advection scheme which is worth exploring further in simulations with realistic protocols which allow meaningful comparisons against observational data. One interesting question is how much the simulations differ with respect to the limiter with or without the mass fixer? The tracers examined here from these simulations are CO₂, CH₄ and CO₂APF. The latter is the anthropogenic CO₂ that is emitted in the simulation period. This is the most challenging tracer for the advection scheme among the three tracers considered here, as it starts from a zero-background state and gradually increases and mixes with the air as the simulation progresses. As we demonstrated earlier, a localized tracer with zero background produces large conservation errors.



(a): LQM-LQM3D with mass fixer on (units ppm) (b): No Fixer – Fixer difference for LQM3D



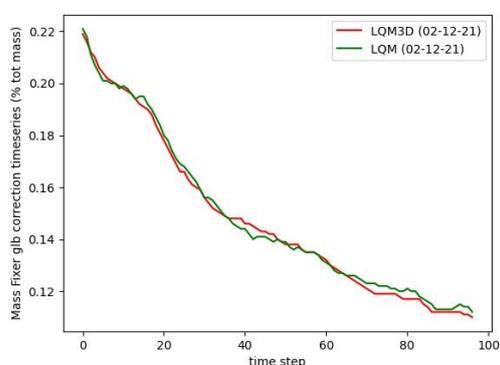
(c): LQM3D-LQM with mass fixer off (d): No Fixer – Fixer difference for LQM

Figure 17: IFS model difference in predicted CO₂APF tracer total column mass in ppm after 1 year of simulation (01-12-2022). Both simulations use COMAD interpolation scheme. Note the difference in plotting scale in (a).

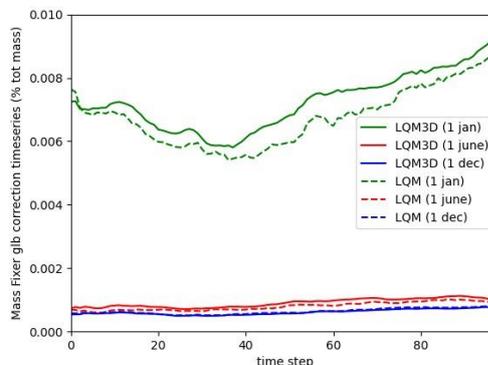
Fig. 17 shows the differences between IFS simulations using the tco399 grid (25 km horizontal spacing) and 137 vertical levels, following the WP7 protocols [Chevallier et al., 2024]. The comparison is made for two different limiters (LQM3D and LQM) on a date corresponding to December 1, 2022 (one year after the simulations start). In both simulations the COMAD scheme is used (an improved version which reduces

CATRINE

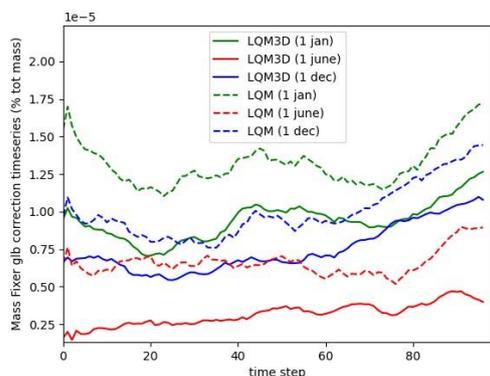
conservation errors and will be described in deliverable D1.2) which reduces conservation error compared with standard cubic-Lagrange interpolation. WP7 CO₂ fluxes have been optimized to correspond to the observed background atmospheric values so this setup allows testing the accuracy of model simulations by comparing to observations. In (a) the difference in the total column of CO₂APF (starting from zero) is shown when the mass fixer is used. We notice that in this case, the difference is small and in absolute value it ranges between 0.0001 to 0.1 in ppm units (column total equivalent). The global mean difference between the two runs is equal to 0.00024 ppm and the Root Mean Square (RMS) norm of their difference is 0.075 ppm. Another feature is that the choice of limiter seems to slightly affect the inter-hemispheric gradient (LQM results in less CO₂ in the SH and more in the NH). When the mass fixer is switched off (Fig. 17c), the two limiters produce larger differences as expected. Both the mean and RMS norm of the difference is approximately 0.21 ppm. Furthermore, when comparing experiments with and without mass fixer even larger differences are seen, in the range 1-1.5 ppm. For LQM3D (Fig. 17b) the mean difference is 1.21 ppm and the RMS norm of the difference: 1.22 ppm and approximately 1.1 ppm for LQM (Fig. 17d).



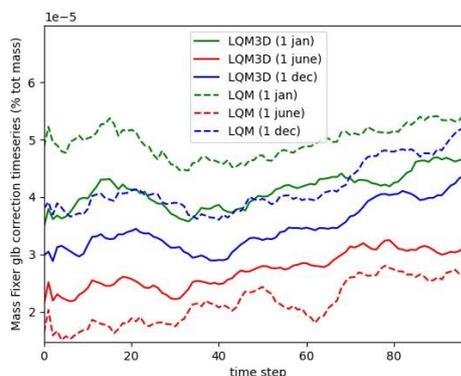
(a): CO₂APF 1 day after initialization



(b): CO₂APF Jan, June, Dec'22.



(c): CO₂ Jan, June, Dec'22.



(d): CH₄ Jan, June, Dec'22.

Figure 18: Time series of mass fixer correction (in percent of the total tracer mass) at each time-step for different tracers from WP7 simulations: CO₂APF (Anthropogenic CO₂ emitted in the simulation period), CO₂ and CH₄.

The above results show that the impact of the mass fixer for CO₂ is overall non-negligible but much smaller than the cases previously examined in this report. The actual correction that the mass fixer computes for different dates for LQM and LQM3D limiter is shown in time series in Fig. 18. The global mass fixer correction per time-step, expressed as a percent of the global tracer mass is plotted there. An interesting feature of these plots is

CATRINE

how much the magnitude of the mass fixer correction for CO₂APF changes over time. For example, in (a) at the second day of the simulation the computed correction is of the order of 0.1-0.2% of the global mass while it constantly reduces dropping to 0.01% after 1 month and to just 0.001% after a further 6 months of simulation. The corresponding correction for the CO₂ and CH₄ fields is two orders of magnitude smaller with much smaller intraseasonal variation. The reason is that the anthropogenic part of the CO₂ (CO₂APF) starts from 0 background value and therefore even after 1 year it is less mixed. Nevertheless, for all considered tracers, the conservation errors the mass fixer corrects are several orders of magnitude smaller than the correction in the experimental cases studied before (compared with Fig. 11). This happens because anthropogenic emissions are much weaker than the more extreme cases studied earlier and because of the very large background values in realistic simulation scenarios such as the one in WP7.

The mass fixer must correct the tracer where large conservation errors occur without affecting the background. This can be achieved by tuning of the so called β parameter in the mass fixer, which controls the strength of the correction according to the strength of the concentration gradients, for details see Diamantakis and Agusti-Panareda [2017], Agusti-Panareda et al. [2017]. The WP7 optimized fluxes and the available observations present an opportunity to tune the value of the β parameter. In Fig. 19 time series from model predicted CO₂ concentrations are compared with daily observations from selected NOAA baseline observatories which have been established to provide data at the most “remote air” of the planet to monitor the true background atmosphere. Some statistics (bias, standard deviation of errors) have been gathered for clarity in Table 1 (left columns). We have compared the performance of two limiters (LQM, LQM3D) without using the mass fixer and with the mass fixer. For LQM3D we tested two configurations of the mass fixer corresponding to $\beta = 1.5$ and 2. Overall, we notice that IFS is close to observations. We also notice the systematic trend to have lower daily means for the CO₂ concentrations when the mass fixer is used (purple, blue, cyan lines). For the Mauna Loa and Samoa stations, simulations with the mass fixer produce more accurate results than simulations without. For the south pole and Alaska station the simulations without mass fixer seem to perform better (notably, the WP7 fluxes were optimized with the LMDZ model which employs a different transport scheme, and this difference could account for a portion of the observed error in these remote regions where transport processes play a dominant role). The Alaska case indicates that the mass fixer may be removing more than it should for the background. Results with $\beta = 2$ are slightly better than those with $\beta = 1.5$ and this suggests that perhaps an even larger value may be needed (larger values will remove less from the background but more from regions with gradients). To investigate further this point using both industrial and remote stations, the daily average XCO₂ (column averaged CO₂ mixing ratio) model time series, from the same experiments, against TCCON have been plotted in Fig. 20. The same statistics (bias, standard deviation of errors) are also gathered in Table 1. Consistently, in all sites the bias is smaller when the mass fixer is used (same is true in other TCCON sites not shown here). This is more noticeable for the simulations with large β and it is also visible in the timeseries (the purple curve is very close to the observations). Therefore, a large β value seems to be a better choice for CO₂ and needs to be further investigated in CATRINE WP2.

Our simulations have also been compared against weekly averaged flask observations for CH₄ in Fig. 21. A difference between CH₄ and CO₂ is that the former is using WP5 flux protocols which have not undergone the rigorous optimization procedure of WP7 and thus the model errors appear larger. An interesting result from these plots is that the mass fixer seems to have a larger impact for CH₄ producing lower concentrations which are closer to observations. This is clearly visible in Fig. 21, and it is also reflected in the lower biases, RMSE and STD (standard deviation of errors) when observations are used as the “truth”. These results also show lower bias and RMSE for LQM3D rather than for LQM when the mass fixer is used while when the fixer is used the differences are more marginal. The

CATRINE

sensitivity with respect to β is also small, which suggests that the optimal value for CO₂ will be suitable for CH₄ as well (however, different β can be used if needed). The standard deviation for LQM with mass fixer is clearly the smallest for the “clean air” locations in Antarctica and Hawaii where we can visually see that the light blue curve (model) is closer to the black dots (observations).

CO2 - INSITU			XCO2 - TCCON		
LQM3D + MF(OFF)			LQM3D + MF(OFF)		
Station	δ	σ	Station	δ	σ
Alaska	0.37	1.51	Karlsruhe	1.31	0.80
Hawaii	0.93	0.64	Xianghe	2.24	1.85
Samoa	0.26	0.46	Izana	0.66	0.41
S Pole	0.04	0.20	Lauder	0.50	0.38
LQM + MF(OFF)			LQM + MF(OFF)		
Station	δ	σ	Station	δ	σ
Alaska	0.65	1.37	Karlsruhe	1.67	0.88
Hawaii	1.25	0.67	Xianghe	2.95	1.97
Samoa	0.38	0.47	Izana	0.97	0.38
S Pole	0.09	0.19	Lauder	0.56	0.40
LQM + MF(ON, $\beta = 1.5$)			LQM + MF(ON, $\beta = 1.5$)		
Station	δ	σ	Station	δ	σ
Alaska	-0.82	1.62	Karlsruhe	0.70	0.76
Hawaii	0.23	0.50	Xianghe	1.73	1.89
Samoa	-0.26	0.41	Izana	0.03	0.43
S Pole	-0.37	0.17	Lauder	0.11	0.42
LQM3D + MF(ON, $\beta = 1.5$)			LQM3D + MF(ON, $\beta = 1.5$)		
Station	δ	σ	Station	δ	σ
Alaska	-0.69	1.57	Karlsruhe	0.59	0.79
Hawaii	0.26	0.58	Xianghe	1.50	1.42
Samoa	-0.16	0.46	Izana	0.01	0.46
S Pole	-0.28	0.16	Lauder	0.16	0.40
LQM3D + MF(ON, $\beta = 2$)			LQM3D + MF(ON, $\beta = 2$)		
Station	δ	σ	Station	δ	σ
Alaska	-0.74	1.57	Karlsruhe	0.55	0.79
Hawaii	0.22	0.57	Xianghe	1.22	1.71
Samoa	-0.13	0.46	Izana	-0.01	0.46
S Pole	-0.23	0.16	Lauder	0.20	0.40

Table 1. Verification statistics from CO₂ simulations. Left: bias (δ) and standard deviation of errors (σ) using insitu observations corresponding to Fig. 19. Right: bias (δ) and standard deviation (σ) against TCCON corresponding to Fig. 20. Units in ppm.

CATRINE

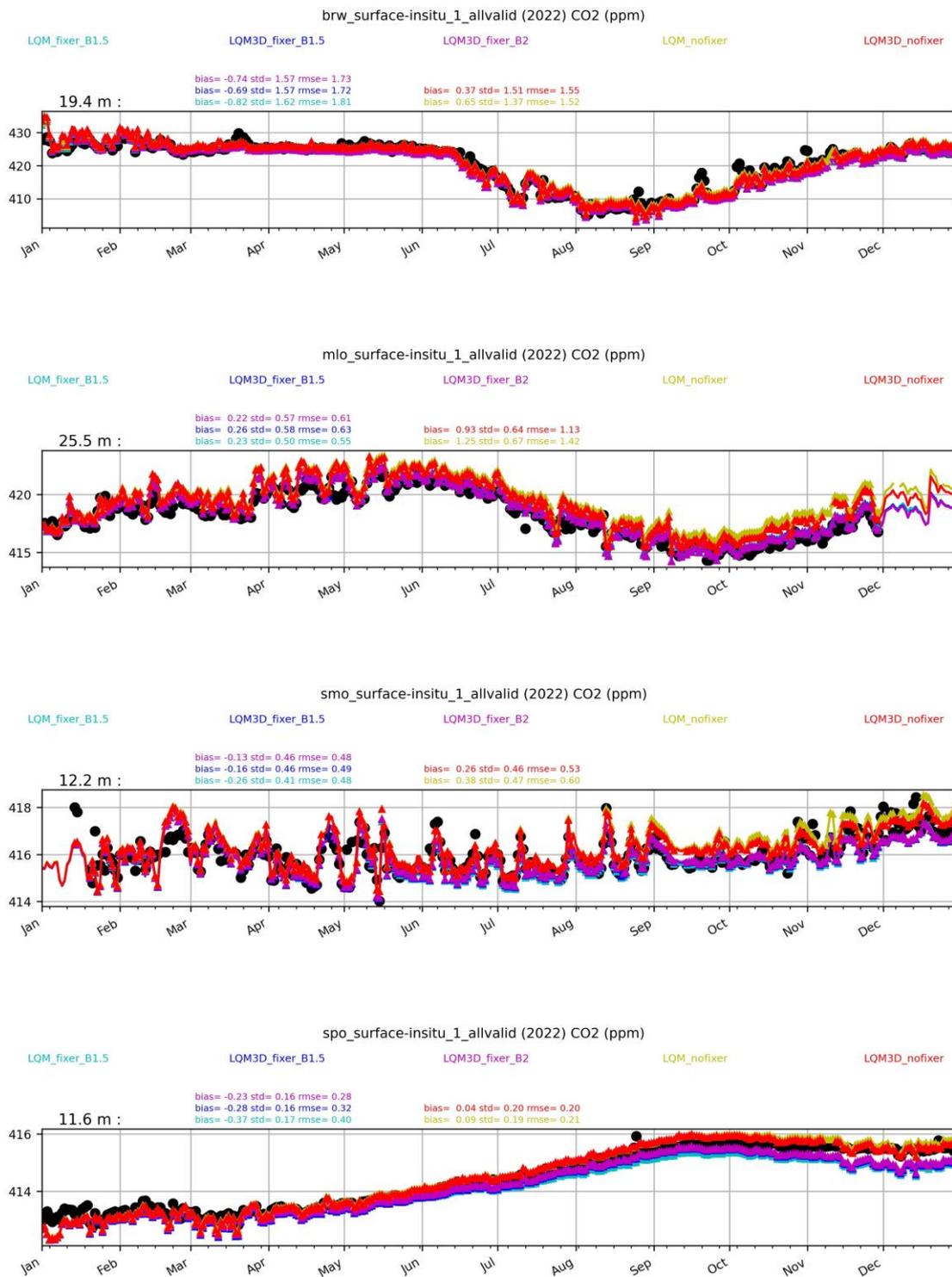


Figure 19. Timeseries (January-December 2022) of daily average CO₂ concentration (ppm) compared against observations at four sites. From the top to the bottom: Barrow station (Alaska), Mt Mauna Loa station (Hawaii), American Samoa station, South pole station.

CATRINE

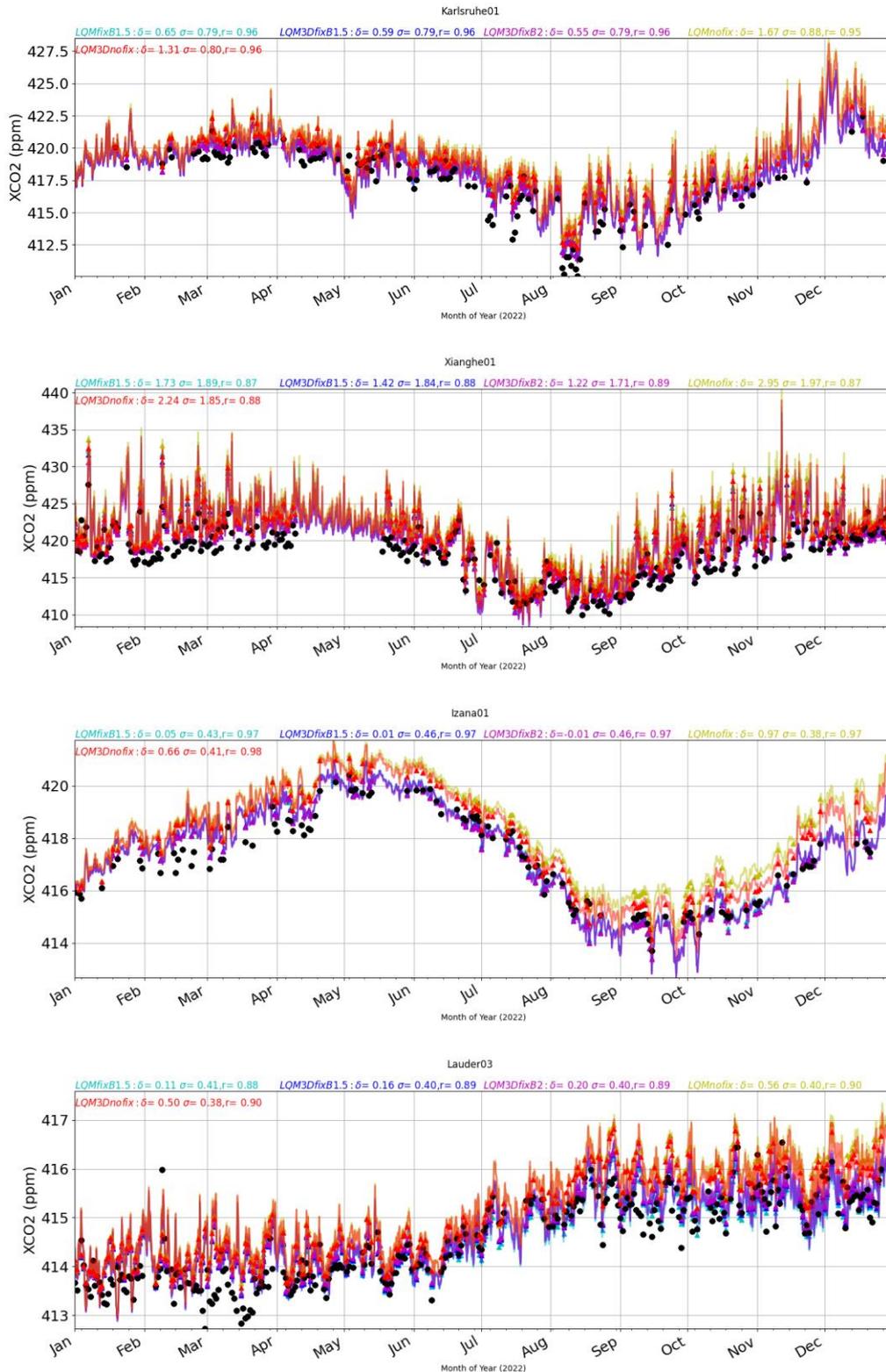


Figure 20. Timeseries (January-December 2022) of daily average XCO2 compared to TCCON observations at four sites. From the top to the bottom: Karlsruhe (Germany), Xianghe (China), Izana (Canary islands), Lauder (New Zealand Southern island). The first two are near areas of industrial/anthropogenic emissions while the last two are away from pollution sites. In each plot the bias δ , standard deviation of errors σ and correlation coefficient is written at the top for each simulation. Units in ppm.

CATRINE

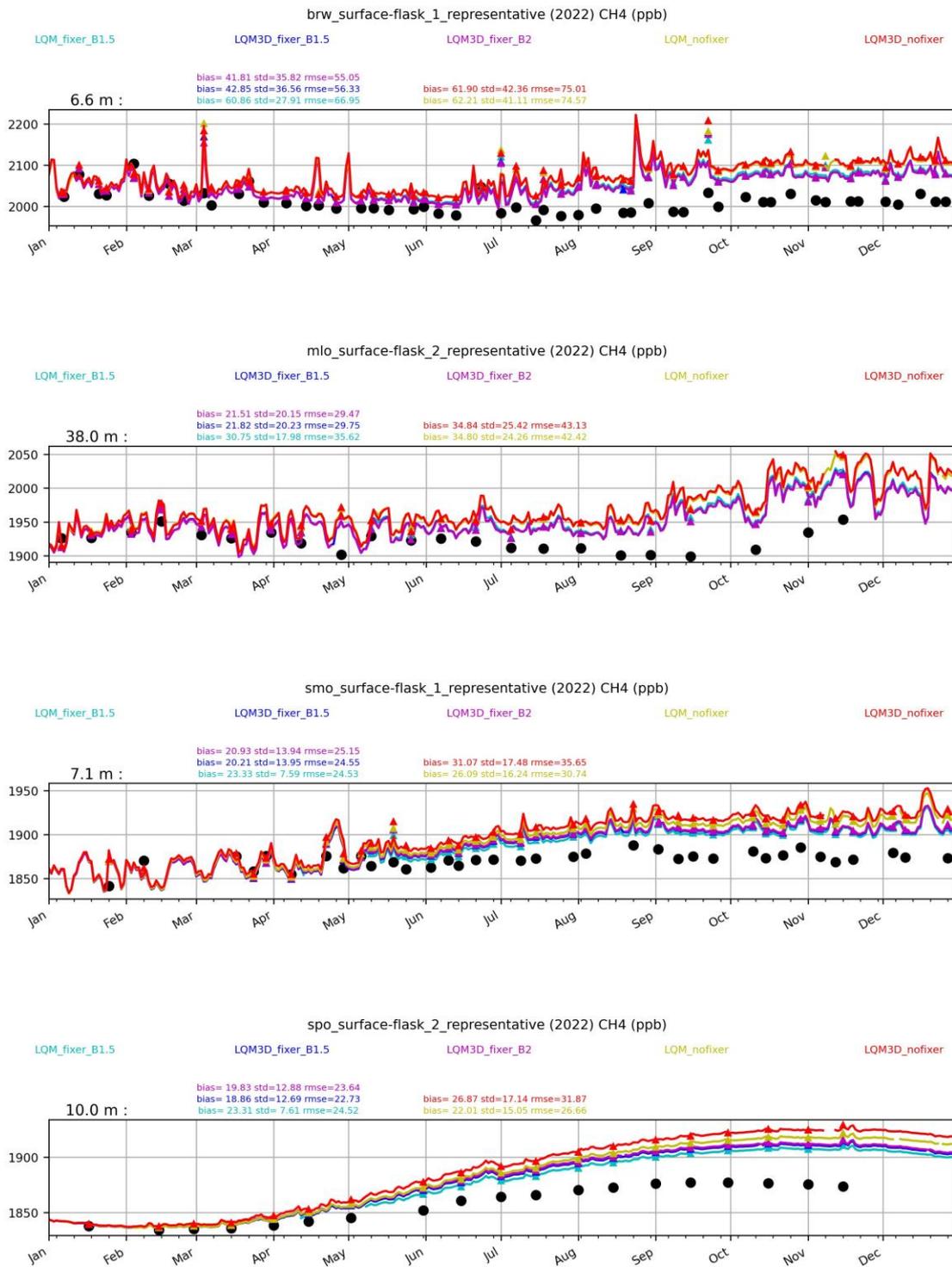


Figure 21. Timeseries (January-December 2022) of weekly average CH₄ concentrations (ppb) compared against flask (weekly averaged observations at four sites. Same locations as in Fig.20: Barrow station, Mt Mauna Loa station, American Samoa, South pole.

7 ICON-ART and comparisons with the IFS

Comparisons between two participating global models in CATRINE, IFS and ICON-ART, have been conducted here. Before comparing the two models, ICON-ART sensitivity experiments were also conducted which helped to identify its best setup for experimentation in terms of stability and accuracy. These experiments employed various combinations of horizontal tracer advection schemes (Miura-type with linear reconstruction, hybrid MIURA3/FFSL with quadratic or cubic reconstruction, both with sub-cycling), horizontal limiters (monotonic Flux-Corrected Transport, positive definite Flux Corrected Transport), and vertical limiters (semi-monotonic reconstruction filter, monotonic reconstruction filter, positive definite Flux Corrected Transport). Among the tested configurations, the hybrid MIURA3/FFSL scheme coupled with monotonic Flux Corrected Transport and monotonic reconstruction filter demonstrated the most favourable performance in terms of mass conservation.

7.1 Tests for ICON-ART transport schemes

The ICON model was implemented in a R3B6L120 configuration, approximating a 25 km horizontal resolution and 120 vertical levels, close to the operational setup used by the German Weather Service and used as one of the main resolutions in other work packages of CATRINE.

To optimize the transport scheme within ICON-ART, a series of sensitivity experiments were conducted, using the plume tracer case study examined in section 5 (rectangular plume tracers), systematically varying horizontal and vertical advection schemes and associated limiters. These experiments aimed to identify a configuration that ensures both stability and accuracy in tracer transport.

transport template	ivadv tracer	itype vlimit	ihadv tracer	itype hlimit
on	3	2	22	3
off	0	0	0	0
stdchem	3	3	22	4
stdchem amip	3	1	22	3
stdchem amip2	3	3	52	4
qv	3	1	52	3
ccmi	3	1	22	3
stdaero	3	2	22	3
hadv52aero	3	2	52	3

Table 2. Transport Template Configurations within ICON. The numbers have the following meaning: vertical advection ivadv tracer (0: no vertical advection; 3: 3rd order piecewise parabolic method); vertical limiter itype vlimit (0: no limiter; 1: 1st order upwind; 2: monotonic reconstruction filter; 3: positive definite flux limiter); horizontal advection ihadv tracer (0: no horizontal advection; 22: 2nd order miura and sub cycling; 52: hybrid method combining the FFSL (Flux-Form Semi-Lagrangian) approach with the Miura scheme); horizontal flux limiter (0: no limiter; 3: monotonic flux limiter; 4: positive definite flux limiter).

The tested transport template configurations are detailed in Table 2. Tracers were initialized in distinct horizontal and vertical source regions (Table 3) to evaluate transport

CATRINE

performance across a range of atmospheric domains. This is as close as possible to the initial setup depicted in Fig. 10a.

	USA	China	Indonesia
Strato (28hPa-32hPa)	38N-43N/80W-74W	35N-40N/110E-116E	10S-5S/110E-116E
UTLS (250hPa-260hPa)	38N-43N/80W-74W	35N-40N/110E-116E	10S-5S/110E-116E
Surface (lowest level)	38N-43N/80W-74W	35N-40N/110E-116E	10S-5S/110E-116E

Table 3. Source regions of the 9 Tracers for the tests of the transport schemes within ICON.

The primary objective of these experiments was to assess the mass conservation properties of each transport scheme. This was achieved by analysing the temporal evolution of total tracer mass. The experimental procedure consisted of:

- **Configuration Setup:** Each transport template (Table 2) was implemented, specifying the chosen advection schemes and limiters. To isolate transport effects, diffusion and other effects besides advection were disabled.
- **Tracer Initialization:** Tracers were initialized with a mass mixing ratio (concentration) of 1 within the source regions defined in Table 3, including the stratosphere, upper troposphere/lower stratosphere (UTLS), and surface levels over the USA, China, and Indonesia. Individual tracer evolution was tracked, with no inter-tracer interactions.
- **Simulation Execution:** Simulations were performed for each configuration to monitor tracer transport over time.
- **Data Analysis:** The temporal evolution of total tracer mass was analysed to evaluate mass conservation. Figs. 22 and 23 present a comparative analysis of two transport templates, *hadv52aero* and *stdaero*, and two source regions, *USA surface* and *Indonesia UTLS* highlighting differences in mass conservation and tracer distribution.

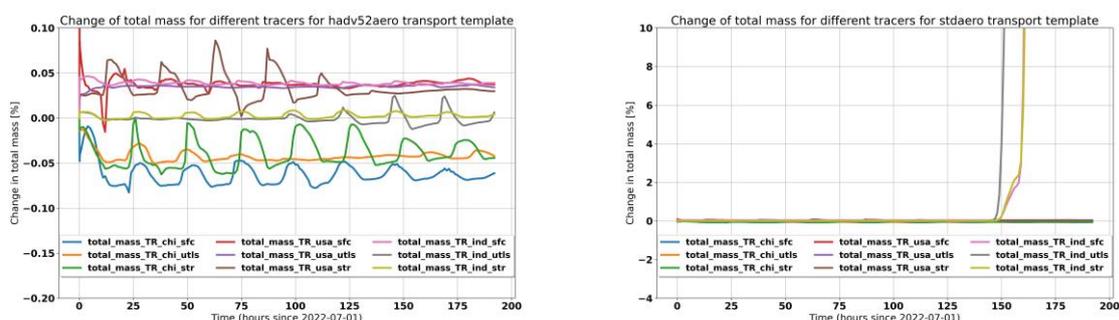


Figure 22. Temporal evolution of total mass for the 9 tracers within ICON for two transport templates. left: *hadv52aero*, right: *stdaero*

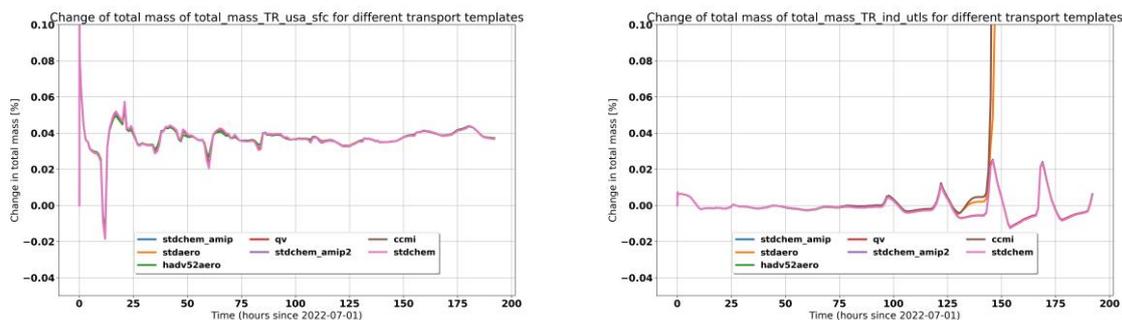


Figure 23. Temporal evolution of total mass for the USA surface tracer (left) and Indonesia UTLS tracer (right) within ICON for the different transport templates.

Further sensitivity tests were performed to examine the impact of configuration variations on tracer evolution. These tests provided insights into the strengths and weaknesses of each scheme, facilitating optimization of the ICON-ART setup. More specifically, the hybrid MIURA3/FFSL scheme, coupled with monotonic Flux-Corrected Transport and a monotonic reconstruction filter (*hadv52aero*), demonstrated superior mass conservation. This configuration was subsequently adopted, significantly enhancing the stability and accuracy of ICON-ART for subsequent studies.

Conversely, *stdaero*, *stdchem amip*, and *ccmi* configurations exhibited mass conservation issues. Specifically, for tracers originating from Indonesia, regardless of initial vertical location, an unexpected and rapid increase in total mass was observed after a few simulation days. This mass increase was accompanied by significant changes in tracer distribution (see Fig. 24 and Fig. 22/23). A common characteristic of these problematic configurations was the use of a second-order MIURA scheme with a monotonic flux limiter. While both components performed well independently, their combination resulted in this anomalous behaviour. The underlying mechanism for this interaction remains unclear and appeared to arise spontaneously.

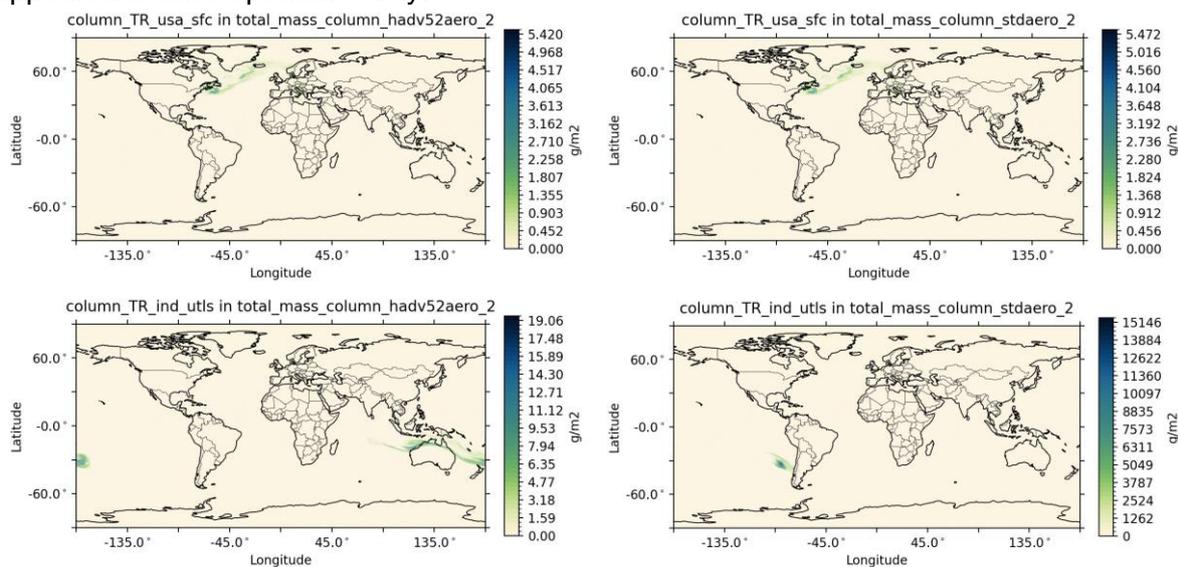


Figure 24. Comparison of plume tracer case study evolution of total tracer columns between ICON transport schemes *hadv52aero* (left) and *stdaero* (right) after 8 days of simulation. In the top row for the surface USA tracer (showing good agreement) and in the bottom row for the Indonesia UTLS tracer (large differences seen there indicating something is going wrong, note the colour scale)

7.1.1 Preliminary results from artificial plume type tracers

Preliminary comparisons between the IFS with the global mass fixer active and inherently conserving ICON-ART (using the best of tested configurations hadv52aero) show broadly similar evolution of passively advected plumes (see Fig. 25). However, differences are expected due to the significant differences between the models, including different forecast winds, grid structures, numerical methods, and parametrizations. Additionally, differences in the vertical coordinate systems and vertical level spacing is another significant factor. It makes it difficult to place the “idealised” plume tracers at the same height and furthermore, because of the different level thickness the same mixing ratio at a specific point may result in slightly different total column masses in the two models. Thus, these comparisons are broadly indicative, primarily aiming to confirm that, despite significant formulation differences and the use of mass fixers to achieve conservation in IFS, the obtained solutions remain comparable. It is worth mentioning that without the mass fixer in IFS much larger differences are present as expected based on our previous results (not shown here). Further tests have been conducted using WP5 and WP7 protocols, which evaluate the advection and overall transport schemes of both modelling systems against observations. These tests will be summarized in detail in Deliverable D5.2, with a sample comparison also provided in the next subsection as part of the ongoing evaluation exercise.

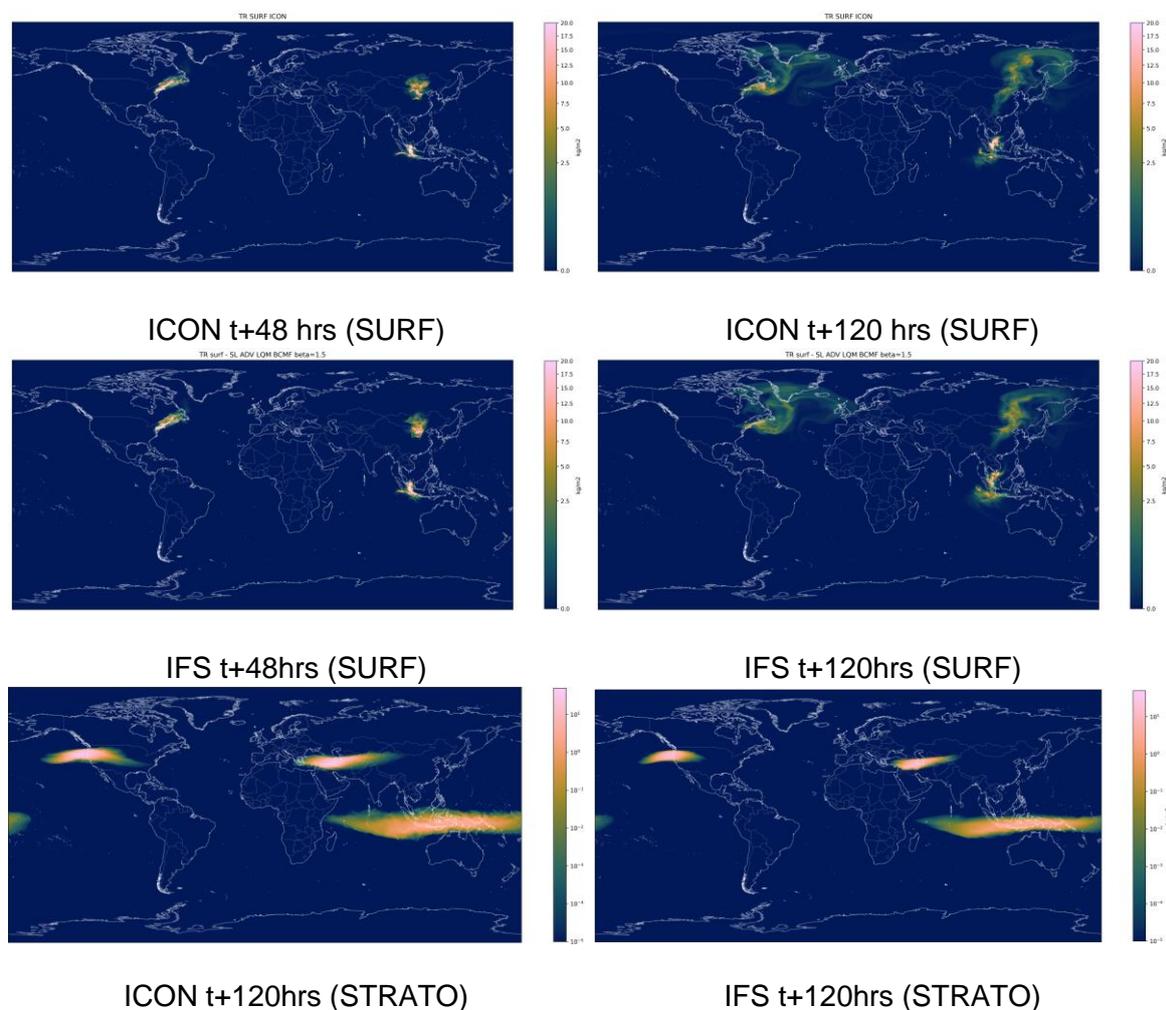


Figure 25. Comparison of plume tracer case study evolution between ICON and IFS for a tracer initialized at the surface (top, middle) and in the stratosphere (bottom) at different lead times. Here the mass fixer is activated in IFS (as in standard runs) to enforce conservation.

7.1.2 Preliminary results from real world tracers

WP5 establishes test-beds for evaluating model performance against observational data. Fig. 26 presents a comparison of CO₂ mixing ratios from the ICON-ART and IFS models with in-situ observations collected during a Dynamics and Chemistry of the Summer Stratosphere (DCOTSS) research flight over the USA. As expected, CO₂ mixing ratios exhibit a strong inverse correlation with flight altitude. For this specific case, both models demonstrate close agreement with the observational data, as quantified by the Taylor diagram in Fig. 26.

The analytical methodologies developed within WP5, including the application of Taylor diagrams, will be leveraged in subsequent analyses within WP1 and WP2 to further investigate model discrepancies and improve understanding of stratospheric processes. The findings within this deliverable will help to understand the findings within WP5/6 and WP7/8.

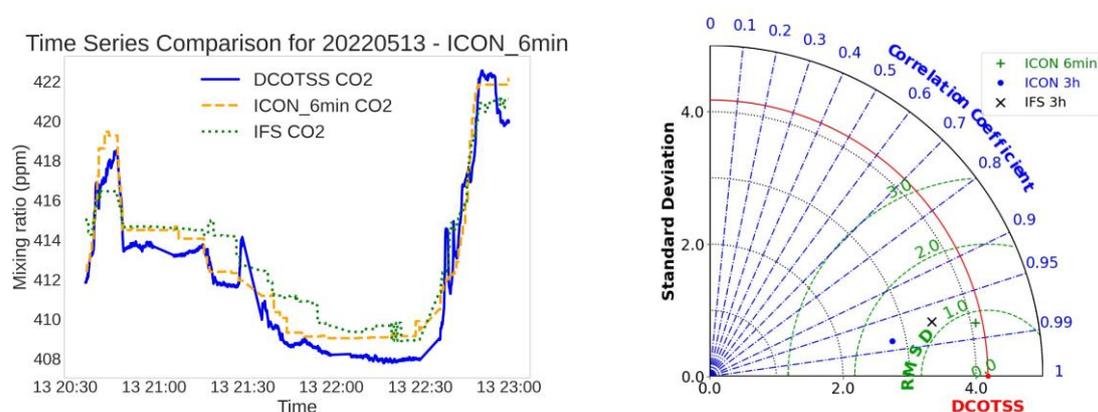


Figure 26. Preliminary comparison of temporal evolution of CO₂ along the flight path of the DCOTSS campaign during May 13th, 2022. Left: Mixing ratio of CO₂. The observational data is shown in blue, ICON-ART in orange and IFS in green. IFS had a 3 hour output step, ICON-ART 6 minutes. Right: Taylor diagram for the flight shown on the left (Note: Here also ICON with 3 hourly output is shown for comparison)

8 Sensitivity in the IFS EDA-based Observing System Simulation Inversion Experiments (OSSE)

An extension of the 4D-VAR IFS assimilation system has recently been developed at ECMWF to jointly optimize the surface emissions of some chemical species (e.g. CO and NO) and greenhouse gases (CH₄, CO₂), alongside their atmospheric concentration and the meteorological state [McNorton et al., 2022]. Since transport represents one of the main sources of uncertainty in emission inversion problems, running sensitivity inversion experiments based on different combinations of mass fixers, limiters and advection schemes can provide an indication of their relative impact on the performance of the system.

8.1 Synthetic emission inversion experiments

An idealised system based on the Ensemble of Data Assimilation (EDA) method has been used to perform sensitivity inversion experiments based on synthetic methane observations. A schematic illustrating this setup is shown in Fig. 27. Akin to a typical EDA, multiple 4D-Var minimizations (10 members) are performed in parallel, each of which assimilates a different set of (perturbed) synthetic observations, while the control doesn't assimilate any observations¹. These observations are created by interpolating the first guess trajectory of the (unperturbed) control member, as if they were satellite CH₄ retrievals of TROPOMI. In this way, the control can be considered as the 'true' state and can be used as a reference to compute the error. The observation error is set to 1.25 ppb for all observations in all experiments. As shown in Fig. 28, the observing system is chosen to be a regularly spaced grid of 25 points surrounding the location of a point source of methane (of magnitude $2.65 \cdot 10^{-7} \text{ kg/m}^2\text{s}$) located at (38N, 118E). A background emission rate (of magnitude $1.53 \cdot 10^{-12} \text{ kg/m}^2\text{s}$) is defined globally, as well as a global initial concentration value of 1000 ppb.

The IFS-based emission inversion system optimizes a global 2D (i.e. single level) species dependent multiplicative scaling factor field, which rescales the value of the emission during each model integration, and it is maintained constant in each 12 hours assimilation window. Each ensemble member is initialized from a perturbed scaling factor field, with perturbations sampled from a prior error covariance matrix B_p with fixed global standard deviation (10%) and horizontal correlation length (300 km). To avoid sampling issues due to the small size of the ensemble, the prior ensemble perturbations are rescaled so that the ensemble spread matches the specified standard deviation value in B_p . The ensemble is used to update the error covariance matrices for meteorological variables and concentrations like in the operational EDA suite. Furthermore, the background methane concentration fields are re-centred around the control at the start of each assimilation window. Lastly, to reduce the impact of non-linearities, all ensemble members are re-initialised in each assimilation window with the control's meteorological fields. The experiments are run for two days (four assimilation cycles). In all of them, three inner loop minimizations are made at spectral truncation resolutions T159, T191, T255, while the final trajectory is run at Tco399 (25km grid spacing).

8.2 Experiment design and results

Table 4 provides a summary of the sensitivity experiments conducted. In these experiments, components of the advection scheme that were found to have a large impact in model simulations are tested. The components tested are labelled next to each experiment identifier and include: COMAD interpolation, two different limiters (LQM and LQM3D), the use of a mass fixer (MF) and different values of its coefficient β which affects the mass fixer correction weight: larger values of β amplify further mass fixer corrections in areas with tracer gradients and reduce them further in the background. In addition, the impact of these components on the tangent-linear (TL) version of the nonlinear trajectory (forward model) and its adjoint (AD) is tested, except for the mass fixer which currently does not have a TL and AD counterpart; it is used only in the trajectory. These will be delivered in WP2, D2.1. The TL model is a key component of the incremental 4D-VAR approach used at ECMWF. It is integrated forward in time and is used to compute the cost (objective) function, which measures the departure of the background model state from the observations. The AD of the TL model performs a backward integration in time within the assimilation window and computes the gradient of the objective function used in the minimization. An accurate TL approximation is derived by a line-by-line translation

¹ Note that we only refer to methane observations here, the assimilation of meteorological observations in both the control and ensemble members is not affected

CATRINE

of the nonlinear model (often quoted as “nonlinear trajectory”) computing first-order linear perturbations with respect to the given model trajectory. There are aspects in the advection scheme such as monotone limiters that are represented by functions that are not fully differentiable. These discontinuities in the TL model can lead to great sensitivity to linear perturbations and even growth of these. To investigate whether full consistency between the nonlinear and the TL model or the smoothness of the linear model has an impact, the exclusion of the limiter(s) from the minimization (and therefore from the TL/AD) was tested. The same was done for the COMAD scheme, which includes a non-differentiable aspect in its trajectory related with the calculation of the perturbed interpolation coefficient needed by the scheme. The current operational version of CAMS uses COMAD only in the trajectory and not in the 4D-VAR minimization. COMAD TL/AD has been recently implemented in IFS to be used in CATRINE and future CAMS operations.

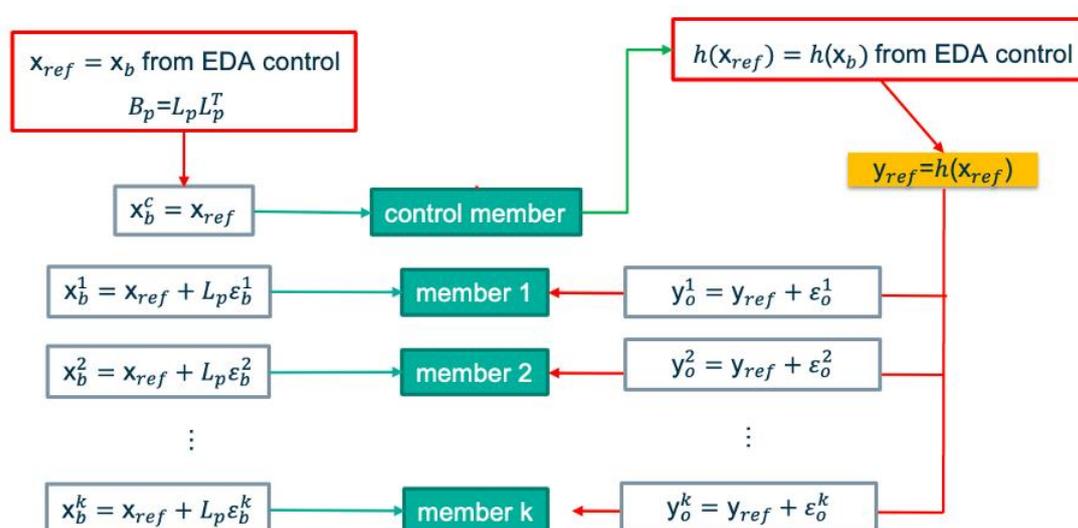


Figure 27. Schematic of the synthetic emission inversion experiments. B_p indicates the emission scaling factor’s prior error covariance matrix, while L_p represents its square root.

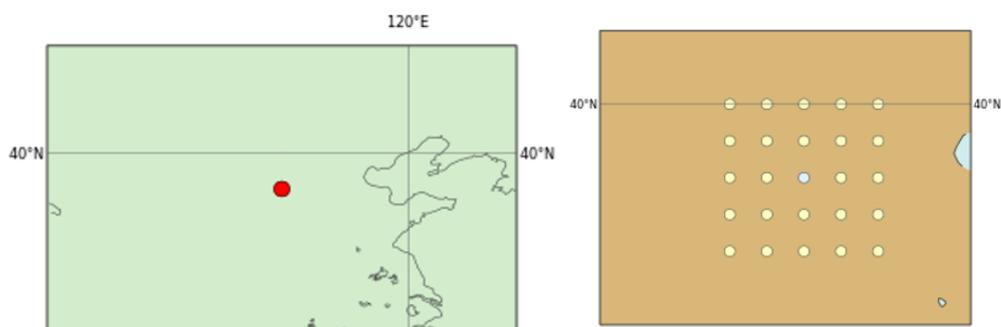


Figure 28. Location of the point source (left) and of the synthetic methane observations (right) in the inversion experiments.

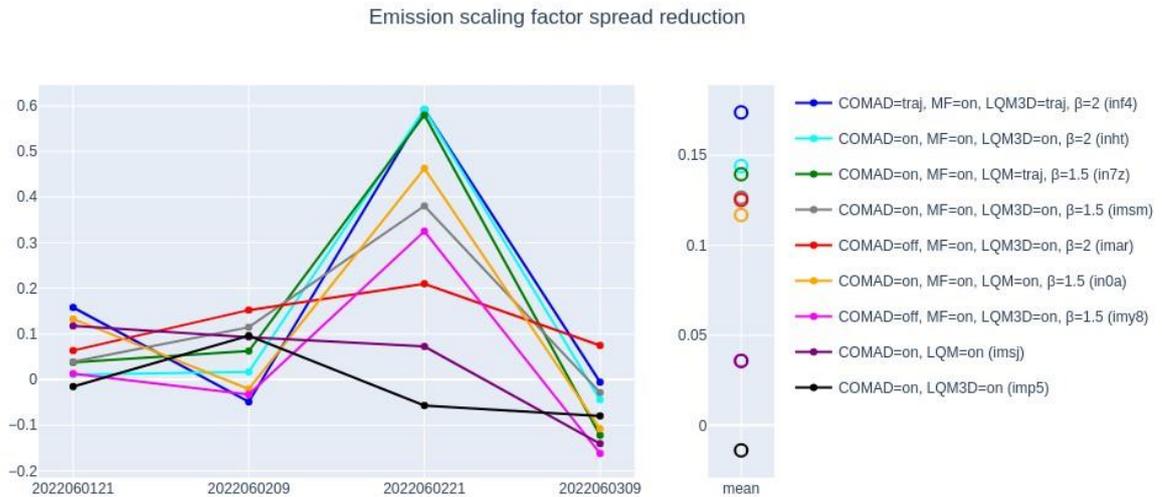


Figure 29. Timeseries of the spread reduction interpolated at the point source location (see Fig. 28) for each experiment listed in Table 4. The mean over the whole period for each experiment is shown on the right. Note that 'traj' in the legend indicates the use of a certain component only in the nonlinear trajectory and the forecast model, excluding it from the minimization.

Two diagnostics were used to assess the sensitivity of the system to different parameters: the reduction in the scaling factor error (er) and the reduction in the scaling factor ensemble spread (sr) at the point-source location. Both are computed as:

$$1 - \frac{\sigma_{pos}}{\sigma_{pri}}, \tag{4}$$

In the case of the error reduction er , σ represents the RMSE of the scaling factor ensemble with respect to the control member before (σ_{pri}) and after (σ_{pos}) the inversion. In the case of the ensemble spread reduction sr , σ indicates the ensemble spread. Note that in both cases, positive (negative) values correspond to a reduction (increase) in the error or ensemble spread, because of the emission inversion.

expid	COMAD	MF	LQM	LQM3D	β
imar	off	on	off	on	2.0
inht	on	on	off	on	2.0
imsm	on	on	off	on	1.5
imy8	off	on	off	on	1.5
in0a	on	on	on	off	1.5
in7z	on	on	on (excl. min)	off	1.5
inf4	on (excl. min)	on	off	on (excl. min)	1.5
imp5	on	off	off	on	N/A
imsj	on	off	on	off	N/A

Table 4. Summary of the sensitivity experiments conducted with the IFS emission inversion system.

CATRINE

Overall, the experiments seem to show some sensitivity in terms of ensemble spread reduction sr to the activation of the mass fixer and the variation of its coefficient β (Fig. 29). The experiments in which the mass fixer is not activated (imp5, imsj) have the smallest overall sr values, while those with mass fixer and larger β values (inht, imar) display larger sr values (hence producing a more *precise* estimate) than the equivalent ones with smaller β values (imy8 and imsm). Sensitivity of the sr reduction to the use of COMAD interpolation is also apparent, since both experiments without the COMAD scheme (imy8, imar) display lower sr values than their COMAD-enabled counterparts (imsm, inht). In contrast, the spread reduction does not seem to be very sensitive to the use of either limiter in presence of a mass fixer. This last result appears to agree with what has been shown in simulations in all previous sections. Lastly, we do observe some sensitivity of sr to the different specifications of the TL/AD setup: experiment inf4 - running without limiters and COMAD in the TL/AD models - displays a larger sr value than its reference (imsm), despite an inconsistency between the model and its adjoint, which is not fully understood at this stage.

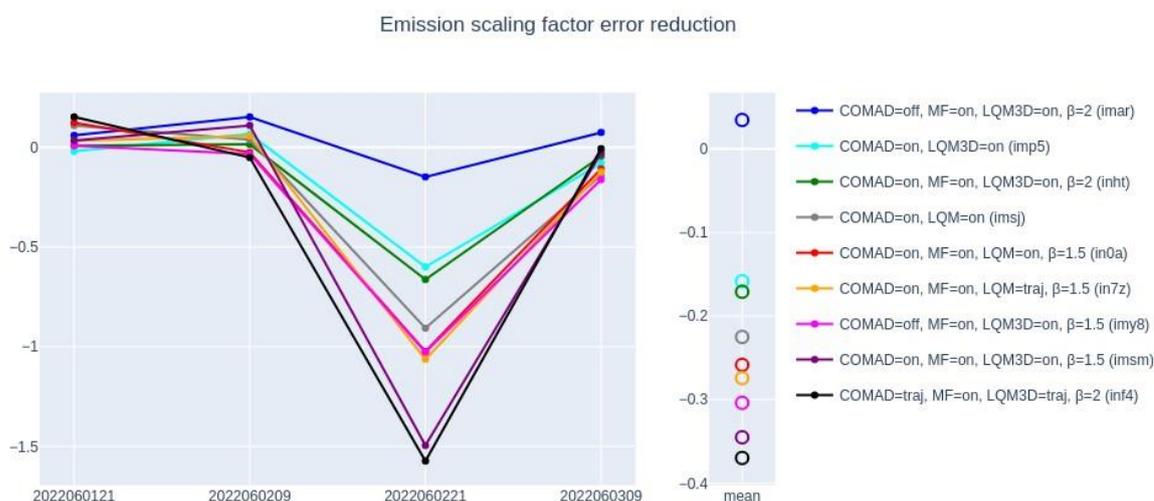


Figure 30. Timeseries of the error reduction interpolated at the point source location (see Fig. 28) for each experiment listed in Table 4. The mean over the whole period for each experiment is shown on the right. Note that 'traj' in the legend indicates the use of a certain component only in the nonlinear trajectory and the forecast model, excluding it from the minimization.

The sensitivity of the error reduction er to the various configurations in Table 4 gives somewhat inconclusive results (Fig. 30). On average, the posterior error increases (i.e. $er < 0$) in all the experiments but imar, as a result of the assimilation, hence producing a less *accurate* estimate of the emission. The mean values are strongly affected by the values in the third cycle, during which a bias develops in all experiments' ensemble (not shown). The reason for this bias is currently unexplained, although it was observed that very large tracer concentrations are present at the point-source location at the beginning of the third assimilation window (not shown), because of reduced turbulent mixing in the boundary layer during night time². The error reduction er seems to be most sensitive to the use of

² The background field of the window starting at 21UTC is the result of a forecast initialised at 2AM local time with a lead time of three hours.

CATRINE

the COMAD scheme and the value of β . The use of COMAD seems to cause an accentuation in the error increase (cf. *inht* vs *imar*), while larger values of β lead both *imar* and *inht* to have a smaller error increase than their counterparts *imy8* and *imsm*.

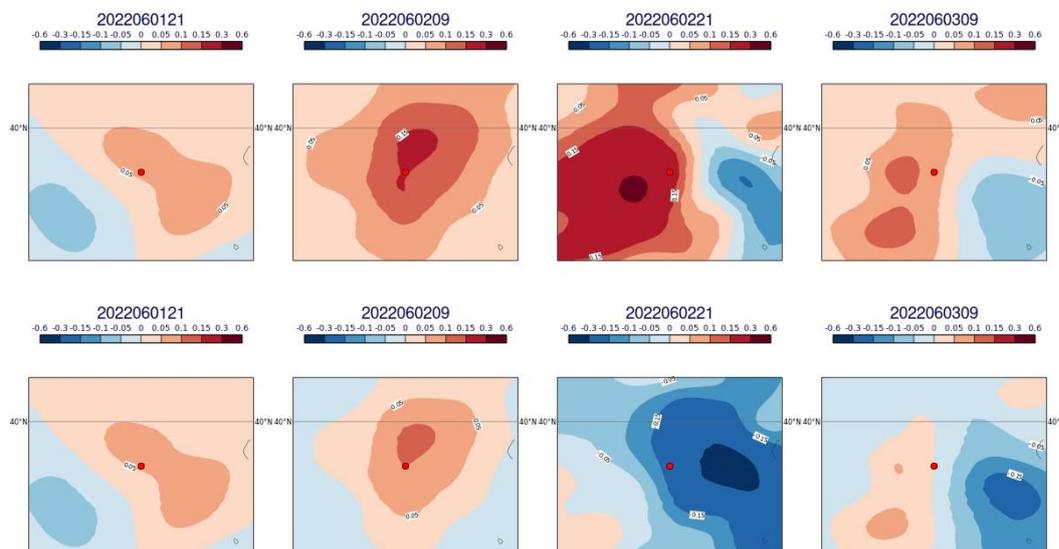


Figure 31. Ensemble spread reduction (top) and error reduction (bottom) in the vicinity of the point source location (red circle) for experiment *imar*. Positive (negative) values indicate a reduction (increase) of the spread after the optimisation.

8.3 Discussion and plans for further experiments

The experiments described in the previous section show sensitivity to some of the components of the advection scheme, but the results cannot be deemed conclusive and further tests and investigations are needed. In particular, it seems that none of the experiments led to a comprehensive reduction in the error in the emission scaling factor over the 2-days period, despite the simplified, idealised setup and its controlled conditions. It is unclear why the large concentration values in the third assimilation window led to a biased estimate of the emission, causing an increase in the error. Overall, most experiments have produced a more precise ($sr > 0$) but less accurate ($er < 0$) estimate of the emission as a result of the assimilation of the synthetic methane retrievals. All in all, the experiment that yielded the best result was *imar* (the only one with $er > 0$ on average), where its spread and error reduction are plotted in Fig. 31.

Further work is needed to understand the current behaviour of the EDA-based OSSE inversion system, especially in view of a new round of sensitivity experiments which are planned in WP2 of CATRINE. It remains unclear at this stage whether the unsatisfactory performance of this system is due to unaccounted sources of nonlinearities, an issue with the way the OSSE has been conceived, or the model and its adjoint itself. Regarding the latter, our current configuration does not include an adjoint for the mass fixer of the advection scheme which could be important for this exercise given the large influence that mass fixer has demonstrated to have in point sources experiments (section 5). The TL/AD counterpart for the mass fixer is currently under development and will be delivered in WP2, therefore we will be able to assess it. Finally, it is worth mentioning that we haven't explored the role and sensitivities of physical parametrizations here such as the

parametrized turbulent mixing scheme which could have a significant impact in model results near the surface. This was beyond the scope of this deliverable.

9 Concluding remarks and future work

In this report, we have conducted a comprehensive analysis and evaluation of the IFS advection scheme, focusing on its conservation properties and accuracy in tracer transport. We have used standard academic test cases, designed new ones, and progressively introduced more complex scenarios. These test cases are highly sensitive to the choice of advection scheme and serve as crucial diagnostic tools by amplifying errors and isolating their individual sources, ultimately facilitating the development of improvements.

Using optimized fluxes from WP7, we assessed the IFS's ability to accurately reproduce observed concentrations and total column quantities of CO₂ in both industrial and clean-air environments. Our results show that the IFS advection scheme can produce accurate simulations. Pre-requisites are the use of the COMAD scheme with cubic interpolation on the cubic Gaussian grid, combined with the IFS tracer mass fixer. Additionally, we evaluated the IFS's sensitivity to various components of the advection scheme. This assessment is a crucial step for developing a system capable of reliably estimating emissions. A detailed list of specific technical conclusions from this study is as follows:

1. Tracers that are well-mixed with air or can be represented as smooth fields produce very small mass conservation and overall transport errors. The largest errors occur where tracers have sharp concentration gradients, changing abruptly like a step function. We distinguish between mass conservation and overall transport error (measured as a norm of the difference of the simulation from the truth and includes aspects such as numerical diffusivity and dispersion) highlighting cases with minimal mass conservation errors but large overall transport errors.
2. A major contributor to mass conservation error is the monotone interpolation limiter (or filter). It significantly impacts results because applying it sacrifices the conservative properties of pure Lagrange polynomial interpolation. However, without the limiter, large undershoots or negative mass values can occur, artificially offsetting positive mass growth. Between the two currently available monotone limiters, LQM3D and LQM, the latter is more mass conserving but also more diffusive. Similarly, linear interpolation strongly damps gradients and often shows very good conservation for tracers away from the boundary and from mountainous terrain. However, it is excessively diffusive for atmospheric composition applications and less accurate than cubic interpolation. The differences between simulations using the two limiters, are considerably reduced when the IFS mass fixer is applied. In real greenhouse simulations, the LQM3D limiter often yields slightly more accurate results.
3. The mass conservation error is largest for a tracer which is initialised or emitted at the terrain following surface level. For the constant pressure levels of the stratosphere, the conservation error is much smaller - typically by more than one order of magnitude. Tests with linear interpolation, which are monotone by design and do not use a limiter, indicate that mountainous terrain enhances mass conservation error at all terrain following levels and this peaks at the near surface level. Furthermore, the lower boundary can act as a mass source when

CATRINE

upward motion occurs and the tracer's vertical gradient is negative, otherwise, it may act as a mass sink.

4. Between horizontal and vertical direction of advection, the latter presents the greatest challenge in terms of conservation errors. This is due to the irregularity of the stretched vertical grid, which leads to larger interpolation and conservation errors, as well as the difficulty to optimally utilize a boundary condition with an advection scheme which is not written in flux formulation. Likewise, the vertical part of the monotone limiter is a major contributor to conservation error.
5. We also find that advection in the meridional direction is more prone to errors than the zonal because of the reduced IFS cubic grid which has slightly non-uniform length and it is slightly mis-aligned in the north-south direction.
6. COMAD interpolation, an IFS option, reduces conservation errors for tracers in the troposphere and especially near the boundary, although significant errors remain. It is especially effective in 'bubble' simulations driven by horizontally converging winds, performing better with the linear Gaussian grid, due to the horizontal wind divergence being better resolved than with a cubic grid. Overall, the combination of COMAD with the cubic grid and cubic interpolation results in smaller mass conservation error growth and therefore it is the recommended choice.
7. No major differences are observed in the performance of the advection scheme or mass conservation properties as the horizontal resolution increases. The mass fixer removes the global mass conservation error from SL advection and remains robust even in extreme test cases with very strong emissions.
8. It is important to emphasize that in our idealised or artificial tracer tests using real weather forecast configuration, advection was the only transport process acting on tracers. This exaggerates mass conservation errors. In practice, parametrized turbulent mixing weakens tracer gradients and reduces SL advection conservation errors. Many of our test cases were specifically designed to stress test the IFS advection scheme. Despite an impressive growth in mass conservation error that some of these cases exhibit, the relative mass conservation error per time step does not exceed 1%, and in realistic simulations is several orders of magnitudes smaller. Therefore, the mass fixer can correct these conservation errors and often improve transport accuracy.
9. Preliminary comparisons with locally and globally conserving ICON model show broadly similar evolution of passively advected plumes, though local differences are present. First results against flight data using WP5 protocols, shows reasonable agreement between the two models. Overall, differences in the two models could be associated not only with the numerical schemes used in the dynamical core of the models but also in differences in vertical resolution (ICON-ART uses fewer vertical levels than IFS), different grids, and different turbulent mixing and convection schemes. A more comprehensive evaluation between the two models will be available by the end of WP5 where accurate flux protocols from WP7 will also be included in ICON simulations.

CATRINE

10. Comparing IFS against CO₂ in-situ and TCCON observations, using WP7 optimized flux protocols, shows a good performance of the IFS model. Sensitivity tests with respect to the limiter indicate that its impact is much smaller in these real-world simulations than in our constructed cases. This is because, in nature, long lived greenhouse gases have very large background concentrations, and they are well mixed with air. Furthermore, anthropogenic emissions are not as strong as the emissions considered in the most extreme of our idealised tracer test cases. Overall, use of mass fixer reduces significantly the bias against TCCON observations both at clean air and industrial sites. Against in-situ observation at clean air sites, its performance is variable: it may increase (South Pole) or decrease the bias (Hawaii) although these results may be affected by compensating errors in the prescribed surface fluxes which have been optimized using a different model (LMDZ). This suggests that tuning the mass fixer could improve further prediction accuracy at clean air sites which will be further investigated in WP2.

11. Finally, sensitivity tests on inversions with a recently made available EDA based idealised OSSE have been conducted revealing sensitivities to the setup such as those found in the forward model. Among these, enforcing conservation with the fixer seems to yield a beneficial impact, although most experiments despite producing more “precise” results (reduced ensemble spread) they fail to produce a comprehensive reduction in the error in the vicinity of point source and they are somewhat inconclusive. The reasons for that are not clear and will be further investigated in WP2. In terms of related specific actions, the priority will be the development of a TL/AD counterpart for the mass fixer, further assessment of OSSE inversion system and understanding better the role and coupling with parametrizations based on investigations which will take place in WP6.

These conclusions strongly suggest that an appropriate setup for CO2MVS should include: (i) the COMAD scheme combined with cubic interpolation; (ii) the cubic-octahedral IFS Gaussian grid; (iii) the use of the IFS tracer mass fixer with the beta parameter set to a value greater than or equal to 2 (to avoid reducing greenhouse gas background values at clean air sites); (iv) the use of an optimized limiter to reduce the conservation error that the mass fixer corrects without introducing too much numerical diffusion. The latter is being explored in Task 1.2 of WP1, while the optimal beta value will be determined with experimentation once all improvements are implemented. The outcomes of this study have informed the design of improvements to the IFS advection scheme, which will be presented in our next deliverable D1.2.

10 References

- A. Agusti-Panareda, M. Diamantakis, V. Bayona, F. Klappenbach, and A. Butz. Improving the inter-hemispheric gradient of total column atmospheric CO₂ and CH₄ in simulations with the ECMWF semi-Lagrangian atmospheric global model. Geoscientific Model Development, 10(1):1–18, 2017. doi: 10.5194/gmd-10-1-2017. URL <https://gmd.copernicus.org/articles/10/1/2017/>.
- T. Becker, T. Rackow, X. Pedruzo, I. Sandu, R. Forbes, M. Diamantakis, P. Bechtold, and I. Polichtchouk. Fixing water and energy budget imbalances in the integrated forecasting system. ECMWF Newsletter, No. 172-Summer 2022, 2022.
- R. Bermejo and J. Conde. A conservative Quasi-Monotone Semi- Lagrangian Scheme. Mon. Weather Rev., 130:423–430, 2002.
- Rodolfo Bermejo and Andrew Staniforth. The conversion of semi-lagrangian advection schemes to quasi-monotone schemes. Monthly Weather Review, 120(11):2622 – 2632, 1992. doi: 10.1175/1520-0493(1992)120<2622:TCOSLA>2.0.CO;2.
- F. Chevallier, A. Agusti-Panareda, M. Krol, W. Peters, and S. Versick. D7.1 design of protocol for preliminary global model intercomparisons. CATRINE deliverable, 2024.
- M. Diamantakis and A. Agusti-Panareda. A positive definite tracer mass fixer for high resolution weather and atmospheric composition forecasts. Technical Report 819, ECMWF, 2017.
- M. Diamantakis and J. Flemming. Global mass fixer algorithms for conservative tracer transport in the ecmwf model. Geosci. Model Dev., 7:965–979, 2014.
- M. Diamantakis and F. Vana. A fast converging and concise algorithm for computing the departure points in semi-Lagrangian weather and climate models. Q.J.R. Meteorol. Soc., 148(743):670–684, 2022.
- S. D. Eastham and D. J. Jacob. Limits on the ability of global eulerian models to resolve intercontinental transport of chemical plumes. Atmospheric Chemistry and Physics, 17(4):2543–2553, 2017. doi: 10.5194/acp-17-2543-2017. URL <https://acp.copernicus.org/articles/17/2543/2017/>.
- ECMWF. *IFS Documentation CY49R1 - Part III: Dynamics and Numerical Procedures*. <https://www.ecmwf.int/en/elibrary/81624-ifs-documentation-cy49r1-part-iii-dynamics-and-numerical-procedures>, 2024.
- M. Hortal. The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ECMWF forecast model. Q.J.R. Meteorol. Soc., 128:1671–1687, 2002.
- S. Malardel and D. Ricard. An alternative cell-averaged departure point reconstruction for pointwise semi-lagrangian transport schemes. Q.J.R. Meteorol. Soc., To Appear, 2015.
- S. Malardel, N. Wedi, W. Deconinck, M. Diamantakis, C. Kuehnlein, G. Mozdzyński, M. Hamrud, and P. Smolarkiewicz. A new grid for the IFS. ECMWF Newsletter, No. 146-Winter 2015-16:30–36, 2016.
- J. McNorton, N. Boussez, A. Agusti-Panareda, G. Balsamo, L. Cantarello, R. Engelen, V. Huijnen, A. Inness, Z. Kipling, M. Parrington, and R. Ribas. Quantification of methane emissions from hotspots and during covid-19 using a global atmospheric inversion. Atmospheric Chemistry and Physics, 22(9):5961–5981, 2022. doi:

CATRINE

10.5194/acp-22-5961-2022. URL <https://acp.copernicus.org/articles/22/5961/2022/>.

F. Rabier, H. Jarvinen, E. Klinker, J.F. Mahfouf, and A. Simmons. The ECMWF operational implementation of four-dimensional variational assimilation. part i: experimental results with simplified physics. 126:1143–1170, 2000.

D.L. Williamson, J.B. Drake, J.J. Hack, R. Jakob, and P.N. Swarztrauber. A standard test set for numerical approximations to the shallow water equations in spherical geometry. J. Comput. Phys., 102:211–224, 1992.

Document History

Version	Author(s)	Date	Changes
0.1 (Initial document created)	Michail Diamantakis and WP1	March 2025	Initial version
0.2	Michail Diamantakis and WP1	20-03-2025	Changes after review by Sylvie Malardel
0.3	Michail Diamantakis and WP1	24-03-2025	Changes after review by Maarten Krol and Anna Agusti-Panareda
1.0	Michail Diamantakis and WP1	25/3/25	Issued version
1.1	Michail Diamantakis and WP1	25/3/25	Minor update to issued version

Internal Review History

Internal Reviewers	Date	Comments
Sylvie Malardel, Maarten Kroll, Anna Agusti-Panareda	20-03-2025	Comments on initial version